A INTELIGÊNCIA ARTIFICIAL DOS *LARGE LANGUAGE MODELS (LLMS)* E OS RISCOS AOS DIREITOS AUTORAIS: DIRETRIZES APLICADAS ÀS PLATAFORMAS E NOVOS DEVERES ÉTICOS-JURÍDICOS PARA SUA UTILIZAÇÃO

Cristiano Colombo[1]
Guilherme Damasio Goulart[2]
DOI: 10.29327/5312711.1-5

Sumário: 1. Introdução. 2. A Inteligência Artificial dos *Large Language Models* (LLM) e os riscos ao Direito de Autor. 3. Diretrizes aplicadas às plataformas e novos deveres éticos-jurídicos para sua utilização. 4. Considerações finais. 5. Referências.

1. Introdução

Nas palavras de Luciano Floridi (2022, p. 11), vive-se "um novo capítulo da história humana", na medida em que a revolução digital traz mudanças que impactam "a maneira como compreendemos nós mesmos, a nossa realidade e a experiência que temos". Cientistas do Direito e da Computação percebem que a orquestração dos campos do saber passa a ser decisiva, visto que novos fatos da vida, causados pelo avanço tecnológico, desde o *input* de dados pessoais aos *outputs* de

1 Pós-Doutor em Direito, Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS). Doutor e Mestre em Direito, Programa de Pós-Graduação em Direito da Universidade Federal do Rio Grande do Sul (UFRGS). Professor Permanente do Mestrado Profissional em Direito da Empresa e dos Negócios da Unisinos; Professor de Graduação em Direito e Relações Internacionais da Unisinos; Professor de Graduação em Direito da Faculdade Verbo Jurídico; Pesquisador Fapergs (2021-2023). E-mail: cristianocolombo@unisinos.br.

2 Doutor e Mestre em Direito pela Universidade Federal do Rio Grande do Sul (UFRGS). Atua como advogado e consultor em Segurança da Informação, Direito da Tecnologia e Proteção de Dados Pessoais. Professor de Direito Civil do Cesuca. E-mail: guilherme@direitodatecnologia.com.

predições algorítmicas de perfis, e, sobretudo, a forma de utilização, exigem um olhar atento e profundo para que os valores sociais subjacentes sejam reconhecidos e preservados.

É compreender a Tridimensionalidade do Direito, na interação entre seus elementos, (1) Fato, (2) Valor e (3) Norma (Reale, 1994, p. 65)[3], bem como considerar que os instantâneos e cíclicos acontecimentos tecnológicos podem vir a ferir valores sociais. A busca em pesquisar soluções que sejam concretas, ainda que provisórias e parciais, em direção à criação de normas, sejam construídas pelos próprios atores, no caminho da autorregulação, ou, heterônomas, apresenta-se como vocação àqueles que percebem a necessidade de tutelar os vulneráveis[4] e salvaguardar o quadro axiológico.

Assim, os Large Language Models (LLMs), ferramentas que utilizam a Inteligência Artificial (IA) e que tem como *input* a linguagem natural, descortinam uma realidade desafiadora, no entanto, de potencial desrespeito dos direitos autorais de criadores, já que as "caixas pretas" dos algoritmos de IA raramente disponibilizam informações sobre quais obras foram utilizadas para o seu treinamento. Do ponto de vista daqueles que já produziram obras, sua criatividade pode ser aprisionada e aprendida pelos processos de treinamento dessas ferramentas, com a reprodução posterior de obras novas baseadas, ou derivadas[5] de suas criações. Um dos aspectos mais problemáticos do uso de ferramentas de IA é justamente as contingências acerca da imposição de regras para o seu funcionamento.

Entre os episódios que merecem destaque, em maio de 2021, Sundar Pichai, CEO da Google, comunicou o lançamento do LaMDA (*Language Model for Dialogue Applications*). Trata-se de

3 "onde quer que haja um fenômeno jurídico, há, sempre e necessariamente, um fato subjacente (fato econômico, geográfico, demográfico, de ordem técnica, etc.); um valor, que confere determinada significação a esse fato, inclinando ou determinando a ação dos homens no sentido de atingir ou preservar certa finalidade ou objetivo; e, finalmente, uma regra ou norma, que representa a relação ou medida que integra um daqueles ao outro" (Reale, 1994, p. 65).

4 A tutela dos vulneráveis aqui se insere no âmbito do uso de obras já criadas para o treinamento das ferramentas de IA. Neste caso, corre-se o risco de que se possa reproduzir indefinidamente exatamente o estilo da obra de determinado autor, "roubando-lhe" o estilo próprio.

5 Cf. art. 5º, inc. VIII, alínea g, da Lei 9.610/98.

um sistema de Inteligência Artificial que "pode conversar com os usuários sobre qualquer assunto" (MIT, 2021).[6] Representa uma tendência no campo de Processamento de Linguagem Natural (NLP), classificado entre os denominados Large Language Models (LLMs) acessível aos consumidores de tecnologia (MIT, 2021). Mais recentemente, experimentamos mais um salto tecnológico, alcançado pelo ChatGPT, atualmente, em sua versão 4.[7]

Quem já utilizou o ChatGPT percebeu que, em alguns casos, ele nega-se a fornecer respostas que potencialmente poderiam ferir alguns princípios éticos. No entanto, uma das características dos dilemas éticos e morais é justamente sua complexidade e dificuldade de aplicação diante de diversos aspectos culturais e sociais. Há o risco, portanto, do algoritmo aprender a se comportar de maneira antiética (Walsh, 2022). Além do mais, tais decisões e contenções são realizadas por seres humanos, que vivem em contextos diferentes, o que pode promover a implementação de vieses e preconceitos.

Nesse sentido, o presente estudo tem como temática proporcionar reflexões sobre os Grandes Modelos de Linguagem (LLMs) e os Direitos Autorais. Têm como problemas de pesquisa: Quais os riscos aos Direitos Autorais, especialmente, em relação àqueles que produziram o conteúdo utilizado como *input* da Inteligência Artificial? Como, no âmbito das plataformas, deve ser construído um quadro adequado, com estabelecimento de novos deveres éticos-jurídicos? Enfim, todas estas questões devem ser sopesadas quando se trata da utilização de LLMs, no contexto da produção de textos a serem produzidos para fins de entrega de trabalhos científicos.

A investigação divide-se em duas partes: na primeira analisar-se-ão os Grandes Modelos de Linguagem e os riscos aos Direitos Autorais; e, na segunda parte, refletir-se-á sobre

6 MIT. A corrida para entender o mundo emocionante e perigoso dos modelos de linguagem para inteligência artificial. MIT Technology Review, 2021. Disponível em: https://mittechreview.com.br/a-corrida-para-entender-o-mundo-emocionante-e-perigoso-dos-modelos-de-linguagem-para-inteligencia-artificial/. Acesso em: 11 ago. 2023.

7 Sobre as atualizações do ChatGPT, é possível acompanhar sua evolução por meio do Release Notes da OpenAI em https://help.openai.com/en/articles/6825453-chatgpt-release-notes.

diretrizes aplicáveis para a construção de novos deveres éticos-jurídicos. Quanto à metodologia, a pesquisa foi teórica, tratando do tema em forma exploratória e descritiva, valendo-se de procedimentos técnicos bibliográficos.

Não se abordará aqui, por uma questão de recorte, a indicação de novos deveres aplicados aos mantenedores das ferramentas de IA que produzem obras sem a participação de seres humanos dando instruções iniciais, ou seja, de obras exclusivamente produzidas por computadores[8]. Também não se tratará sobre a proteção de dados pessoais, tanto nas situações destes dados serem tratados para as operações de treinamento da LLM, tampouco dos dados recolhidos nas interações de seres humanos com essas tecnologias.

2. A Inteligência Artificial dos *Large Language Models (LLMs)* e os riscos ao direito de autor

2.1. A Inteligência Artificial dos *Large Language Models (LLMs)*

Em artigo seminal sobre os riscos dos LLMs, denominado "*On the Dangers of Stochastic*[9] *Parrots: Can Language Models*

8 Sobre o tema ver Bürger, M. L. F. de M. and Corrêa, R. (2020) 'Dos pincéis aos algoritmos: a titularidade das expressões artísticas e criativas resultantes da aplicação da inteligência artificial' in Erhardt Júnior, M., Catalan, M. and Malheiros, P. (coord.) *Direito Civil e Tecnologia*. Belo Horizonte: Fórum. Ver também Parlamento Europeu. *Resolução do Parlamento Europeu, de 20 de outubro de 2020, sobre os direitos de propriedade intelectual para o desenvolvimento de tecnologias ligadas à inteligência artificial* [online]. Available at: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0277_PT.pdf (Accessed: 14 August 2023, p. 7: "Sublinha a diferença existente entre as criações humanas assistidas pela IA e as criações geradas pela IA, sendo que estas últimas colocam desafios em matéria de proteção dos DPI, suscitando nomeadamente questões relacionadas com a titularidade da propriedade ou da invenção e com a remuneração adequada, bem como questões relacionadas com a potencial concentração de mercado; considera, além disso, que cabe fazer a distinção entre os DPI para o desenvolvimento de tecnologias de IA e os DPI eventualmente concedidos às criações geradas pela IA; salienta que o atual quadro dos direitos de autor continua a ser aplicável quando a IA é apenas utilizada por um autor como ferramenta de apoio no âmbito de um processo de criação".

9 De acordo com Carvalho (2021, p. 506): "Nesse sentido, autores da ciência da computação diferenciam o caráter determinístico ou estocástico dos ambientes virtuais. O ambiente será determinístico quando o estado seguinte dos objetos analisados puder ser completamente determinado pela previsão dos efeitos de

Be Too Big?", publicado no ano de 2018, lideradas por Emily Bender e Timnit Gebru, foram apresentadas importantes reflexões, alertando sobre possíveis danos ligados à utilização (Bender, Gebru, McMillan-Major and Shmitchell, 2021). Seus estudos passaram pela análise histórica, destacando que os primeiros modelos de linguagem datam do início da década de 80, voltados, primordialmente, ao *Automatic Speech Recognition* (ASR), ou seja, transformar a fala humana em texto escrito, bem como a *Machine Translation* (MT), tendo como tarefa a tradução de textos para outras línguas (Bender, Gebru, McMillan-Major and Shmitchell, 2021). O grupo de estudiosos também apresentou a definição dos LLMs, no sentido de que são "treinados em tarefas de predição de sequência de caracteres", indicando "possíveis símbolos (caracteres, letras ou mesmo uma sequência, *string*)", tomando como base um "contexto anterior", um "texto como entrada", de forma "não supervisionada", tendo como resultado "pontuações ou sequências de caracteres" (Bender, Gebru, McMillan-Major and Shmitchell, 2021). Interessante salientar que, ao longo dos anos, podem ser apontadas diversas iniciativas de LLMs, como o ERNIE, com suas aplicações em chinês, o BERT, na língua inglesa, a NVIDIA, treinando modelos com dados da Wikipédia, o Switch-C, na estrutura decisional, até as recentíssimas versões do ChatGPT, desenvolvidas pela OpenAI (Bender, Gebru, McMillan-Major and Shmitchell, 2021).

Para entender realmente o problema, é preciso analisar – mesmo que de forma resumida e simplificada – a forma com que as ferramentas de LLM funcionam. O primeiro passo é o treinamento, que nestas situações ocorre com o consumo dos mais variados tipos de dados, que passam desde obras acadêmicas e de literatura, informações factuais e até mesmo interações entre usuários ocorridas em redes sociais ou fóruns de discussão. Esse treinamento visa construir um modelo que consiga prever a possibilidade da ocorrência das

determinada ação sobre o estado atual. Por outro lado, será estocástico quando o ambiente for apenas parcialmente observável, ou seja, quando seus estados seguintes não puderem ser simulados de forma fidedigna. Ocorre que grande parte das situações concretas são tão complexas que é impossível dar conta de todos os seus aspectos não observáveis, razão pela qual são tratadas como estocásticas".

palavras dentro de certos contextos. Isso não ocorre da maneira tradicionalmente feita em programas de computador, em que o programador estabelece previamente os passos que o algoritmo deve seguir. O programador humano não analisa e prevê antecipadamente todas as possibilidades de formação de textos diante dos *inputs* dados pelos usuários e cria todas as situações de atuação (o que seria impossível diante do caráter aleatório e tendente ao infinito de associações entre palavras). Ao contrário, por meio do uso de *machine learning* e redes neurais (Lee and Trott, 2023), o algoritmo apreende contextos e tenta prever qual a palavra mais adequada e provável que virá a seguir. A forma com que tais algoritmos conseguem predizer as palavras passa também por sua classificação. Cada palavra é classificada em vetores, que são informações numéricas, que dependem não somente do seu sentido, mas também dos possíveis contextos em que ela é usada. Tomemos como exemplo a palavra gato (Lee and Trott, 2023)[10]: a depender do contexto em que ela é usada, existirá um vetor diferente que marcará a sua pertinência naquela situação. Vetores diferentes marcam a mesma palavra em "O gato é um animal de estimação" e "Fulano adotou um gato". Não se trata somente de vetores diferentes para funções sintáticas distintas das palavras, mas vetores para contextos, envolvendo uma análise pragmática. Isso faz com que, virtualmente, a mesma palavra possa ter vetores ilimitados, que marcam contextos que são apreendidos com base nos dados de treinamento. Por isso que tais modelos precisam consumir grandes volumes de informações para que consiga marcar os contextos em que as palavras são usadas e que se torne, assim, assertivo e eficiente. Portanto, "cada vetor representa um ponto em um 'espaço de palavras' imaginário, e as palavras com sentidos similares são colocadas mais próximas" (Lee and Trott, 2023)[11]. É possível dizer, neste sentido, que podem existir vetores que aproximam "Porto Alegre" de "Curitiba"[12], enquanto capitais, sem perder de vista que o mo-

10 Usamos e desenvolvemos a mesma palavra usada pelos autores neste ensaio.

11 Segundo Lee e Trott (2023), diante da complexidade da tarefa, podem existir centenas ou milhares de vetores de espaço, com múltiplas dimensões de aproximações. O GPT versão 3, por exemplo, usa vetores com até 12.288 dimensões

12 Os autores Lee e Trott usaram, é claro, nomes de cidades diferentes.

delo ainda precisa estar pronto para lidar com interações com a expressão "Porto Alegre" não como cidade ou capital, mas como um porto com a característica de ser alegre (o que demandaria vetores diferentes, diante do contexto).

Além disso, o modelo é organizado por meio de camadas (*layers*). Essas camadas buscam ajustar o contexto em que as palavras-vetores são usadas, sendo que as expressões vão passando por diversos *layers*. No exemplo dado por Lee e Trott (2023)[13], uma frase contendo a palavra "banco" vai passando pelos *layers* para que ele possa verificar a função desta palavra (diferenciando, por exemplo, o banco onde alguém pode se sentar, de um banco – instituição financeira).

Ocorre que os LLMs, por lidarem com a linguagem humana, podem gerar sérios problemas como a própria exclusão de utilizadores não falantes da língua inglesa, dada a supremacia dos *inputs* em seus modelos. Nessa linha, discriminações étnicas, de gênero, etarismo, e, também, voltadas às pessoas com deficiência, foram identificadas como questões extremamente sensíveis, na sua utilização. A partir de testes realizados, foram identificados inúmeros vieses como associações estereotipadas para determinados grupos, nos *outputs*. Um exemplo bastante surpreendente é que um determinado LLM estabeleceu links entre pessoas com deficiências e palavras negativas (Bender, Gebru, McMillan-Major and Shmitchell, 2021).

Também se reconhece que não há técnicas confiáveis para controlar os comportamentos dos LLMs. Apesar de existirem meios de moderação e de *reinforcement learning* que produzem resultados, eles não são perfeitamente efetivos (Bowman, 2023, p. 5)

Cumpre destacar que tais fatos justificam a metáfora dos *Stochastic Parrots* ou "Papagaios Estocásticos", como *headline* da pesquisa realizada por Bender e Gebru, cotejando os LLMs com o comportamentos destas aves, que imitam as pala-

13 Os autores informam que a versão 3 do GPT possui 96 layers e que: "[...] *that the first few layers focus on understanding the sentence's syntax and resolving ambiguities like we've shown above. Later layers* [...] *work to develop a high-level understanding of the passage as a whole*". Cada um dos layers realiza 96 operações de atenção, para tentar prever a próxima palavra, o que implica em 9.216 operações cada vez que busca realizar tal previsão.

vras sensorialmente coletadas, sem compreenderem os seus significados, ou mesmo, sem terem a capacidade de uma reflexão ética sobre os sons que emitem. Não conseguem perceber os potenciais danos e situações de constrangimento que podem ser geradas.[14] Neste sentido, coloca-se em xeque a fiabilidade dos modelos de linguagem, inclusive, na medida em que situações que não aconteceram já foram apresentadas como se verdadeiras fossem, como a condenação de um advogado estadunidense por citar artigos de leis e jurisprudências falsas, ao peticionar, por meio de ChatGPT (MIT, 2023b). Sendo assim, grandes modelos de linguagem (LLM) podem espalhar a desinformação, bem como serem carregados de conteúdos abusivos (MIT, 2023a), inverdades, merecendo toda a atenção.

No entanto, mesmo diante deste cenário negativo, em recentíssimo estudo, publicado no mês de agosto de 2023, um ponto muito positivo foi apontado para os LLMs, mais precisamente, pela observação dos resultados do ChatGPT. Concluiu-se que houve significativa redução de "discrepâncias de habilidade de escrita", no ambiente organizacional (MIT, 2023a). Empregados "menos experientes", por meio da utilização da ferramenta, aproximaram-se da qualidade dos textos daqueles mais "talentosos", minimizando assimetrias (MIT, 2023a). Observe-se que, além disso, aqueles que se valeram do ChatGPT "levaram 40% menos tempo para completar suas tarefas" e "receberam notas 18% mais altas" (MIT, 2023a). Dessa forma, desarrazoado parece ser "revogar" ou "banir" o ChatGPT das vidas das pessoas e do fluxo do cumprimento de tarefas no âmbito das organizações. Compreende-se ser impraticável cessar a sua utilização.

Por outro lado, outro campo extremamente sensível de observação são potenciais riscos aos direitos autorais, já que os LLMs tomam como *input* textos que são classificados como obras intelectuais. É hora de refletir, cumprindo, àqueles que se dedicam às interações entre as novas tecnologias e o Direito, buscar caminhos para a continuidade da tecnologia, com o

14 Putini, J. (2010) 'Papagaios são isolados em zoológico após ofenderem vários visitantes', *R7,* Hora 7, [online]. Available at: https://noticias.r7.com/hora-7/fotos/papagaios-sao-isolados-em-zoologico-apos-ofenderem-varios-visitantes-02102020 (Accessed: 12 August 2023).

afastamento ou, no mínimo, mitigação de danos aos criadores das obras de espírito.

2.2. Os Riscos a Direitos Autorais no uso de LLMs

Observando o funcionamento dos LLMs, que é membro da família da Inteligência Artificial, alimentam-se eles de conteúdo escrito preexistente (*inputs)*, aplicando instruções algorítmicas (reconhecendo padrões sobre caracteres, palavras, inclusive, identificando sua função sintática, na busca em predizer respostas), para, ao final, como resultado (*output*), oferecer uma nova sequência de texto ou diálogo. É possível afirmar que, em todas as fases, existem vivas interações com a temática dos Direitos Autorais, apresentando possíveis riscos de violações a direitos.

Como *input*, a Inteligência Artificial dos LLMs consome "uma grande quantidade de trabalhos criativos e dados como parte de seu processo de aprendizado" para se produzirem textos "como os humanos" (Bonadio, Dinov and McDonagh, 2022). Os escritos que alimentam a Inteligência Artificial podem vir a ser perfeitamente classificados como "obra intelectual", por serem "textos", que estão expressamente previstos no artigo 7º, da LDA. Portanto, trata-se de um gatilho para a proteção de toda ordem de direitos de autor, sejam morais ou patrimoniais.

Nesse sentido, o artigo 24 da Lei 9.610/98, que dispõe sobre o rol de direitos morais, estabelece a íntima ligação entre a pessoa humana do criador e o que foi criado. Aqui se destacam os direitos de "reivindicar, a qualquer tempo, a autoria da obra", "ter seu nome, pseudônimo ou sinal convencional indicado ou anunciado, como sendo o do autor, na utilização de sua obra", bem como "modificar a obra, antes ou depois de utilizada". Dessa forma, aqueles que se servem de conceitos, de expressões, de frases redigidas por alguém e que as utilizam, no todo ou em parte, em outros textos, tem o dever jurídico de promover a citação direta, se *ipsis litteris*, ou, indireta, caso preservem seus elementos. Assim, há riscos a direitos autorais na utilização de obras literárias sem

a referência a base de dados do autor do *input,* em LLMs, a ser enquadrado como plágio. Segundo o Manual de Direitos Autorais do Tribunal de Contas da União: "O plágio é a violação ao direito autoral em que a obra alheia é apresentada como própria, ainda que de forma "disfarçada" (TCU, 2020). No caso do ChatGPT, são formadas novas frases, todavia, valendo-se de textos preexistentes. Na medida em que alguém se diz autor, sem sê-lo, tem-se plágio.

Perguntando ao ChatGPT, como prompt "Se eu utilizar um texto gerado por você, chatgpt, e não fizer referência, isso pode ser considerado plágio?". A resposta foi:

> Sim, utilizar um texto gerado pelo ChatGPT sem dar créditos ou fazer referência a ele pode ser considerado plágio, especialmente se você apresentar o texto como se fosse seu próprio trabalho original. Plágio envolve a utilização não autorizada ou não atribuída do trabalho intelectual de outra pessoa, incluindo textos, ideias, e outros tipos de criações. Ao utilizar um texto gerado pelo ChatGPT, é importante fornecer atribuição adequada, indicando que o texto foi gerado por uma inteligência artificial, como "Texto gerado por ChatGPT, um modelo de linguagem desenvolvido pela OpenAI". Isso demonstra transparência e ética na utilização de recursos de IA e evita qualquer mal-entendido quanto à autoria do texto (CHATGPT-3.5, 2023).

Outrossim, ressalte-se, que os direitos morais são inalienáveis e irrenunciáveis, não admitindo, portanto, que, por cessão, se transfira a condição de criador para terceiro, como preceitua o artigo 27 da LDA.

Tendo fornecido o seguinte prompt: "Por que você não faz referência quando responde aos questionamentos?", obteve-se a seguinte resposta:

> Eu sou um modelo de linguagem desenvolvido pela OpenAI, e as respostas que forneço são geradas com base em um vasto conjunto de dados de treinamento. Minhas respostas não são retiradas de fontes específicas, mas sim construídas com base no conhecimento geral que adquiri durante o treinamento. No entanto, é importante observar que, mesmo que eu não faça referência explícita a fontes específicas, meu treinamento é uma composição de informações de diversas fontes disponíveis até a minha data de corte em setembro de 2021. Portanto, ao oferecer informações, estou

> essencialmente representando o conhecimento acumulado até aquele ponto, e não estou reproduzindo texto diretamente de uma fonte particular. Quando se trata de ética e credibilidade, é sempre recomendável verificar as informações que você obtém de várias fontes confiáveis e dar créditos apropriados se estiver utilizando informações de fontes específicas em seus trabalhos (CHATGPT-3.5, 2023).

Assim, na versão GPT3.5, de livre acesso, verifica-se que não há o apontamento de fonte específica, sendo indicado um "conhecimento acumulado", com data de corte de setembro de 2021, o que não se aproxima do que determina a legislação.

No que toca aos direitos patrimoniais de autor, adverte Bonadio, Dinov e MacDonagh, que a utilização de "recursos materiais" e "processamento de dados como *input*" poderá violar o "direito de reprodução" (Bonadio, Dinov and McDonagh, 2022) dos criadores. Nessa linha, os LLMs trazem riscos aos direitos patrimoniais do autor, que são o de "utilizar, fruir e dispor da obra literária, artística ou científica", com exclusividade, na forma do artigo 28 da Lei dos Direitos Autorais. Pensemos na utilização de toda a criação, ou de forma parcial, como *input*, para a escrita de um novo livro, por meio de LLM. Na tradução de uma obra na íntegra para outra língua, como do inglês para o português, valendo-se de LLM de Machine Translation (MT), ou, ainda, na adaptação do texto de uma obra de ficção científica para histórias em quadrinhos (HQ). Em todos estes casos, a ausência da "autorização prévia e expressa do autor", nos termos do artigo 29, da mesma lei, para reproduzir parcialmente, adaptar e traduzir, violaria direitos autorais. No caso, não se trata de "pequeno trecho", não sendo o suficiente a mera citação de quem foi o autor. Ainda, enquanto no *input* pode ser vista a obra originária, a "criação primígena", nos termos do artigo 5º, inc. VIII, alínea "f" da LDA, o *output* pode ser potencialmente qualificado como uma "obra derivada". Dessa forma, mais um risco aos direitos autorais.

Saliente-se que, nos termos do artigo 46, VIII, da Lei 9.610 de 1998, na hipótese de utilização de "pequenos trechos de obras preexistentes", "sempre que a reprodução em si não seja o objetivo principal da obra nova" e, ainda, "que não prejudique a exploração normal da obra reproduzida nem

cause um prejuízo injustificado aos legítimos interesses dos autores", é possível a utilização sem "a autorização expressa e prévia", no entanto, o descumprimento do dever de citação, seja direta ou indireta, conforme couber, também opera riscos ao direito moral de autor.

Quanto às instruções algorítmicas, importa destacar que, por serem linguagem de programação, também têm proteção, a teor do artigo 7º, XII da LDA, bem como da própria Lei do Software, Lei 9.609/1998, com seus direitos ali descritos.

No tocante aos *outputs*, nos Grandes Modelos de Linguagem, sustentam Bonadio, Dinov e MacDonagh que os "produtos finais", podem vir a ser considerados "adaptação a trabalhos preexistentes", e, que "qualquer constatação de infração dependerá se os elementos preexistentes possam ser reconhecidos no produto final" (Bonadio, Dinov and McDonagh, 2022). Trata-se de importante contribuição para a avaliação de riscos a direitos autorais, no sentido de que, além da via de pura e simples comprovação de que o texto primígeno foi utilizado na obra gerada pela IA, por meio da aferição de seu *input*, também se poderá realizar a comprovação de plágio a textos preexistentes, quando demonstrados que estes se projetam substancialmente, ainda que em parte, e, possam vir a ser reconhecidos no produto final. Essa projeção, ou uso não substancial de uma obra protegida, pode ser comparada com o que se conhece na doutrina americana como "*non-expressive reading*", no âmbito do *fair use*, ou seja, o princípio que se aplica nos casos em "que há algo a ser aprendido sobre uma obra protegida por direitos autorais, além de sua expressiva autoria" (Grimmelmann, 2016, p. 661).[15]

15 Destacamos que a citação do "*non-expressive readingg*" deve levar em consideração que o sistema de *common law* americano tende a ser mais aberto às resoluções de novos problemas do que os sistemas de *civil law* em face da importância que se dá aos precedentes. Assim, seria mais fácil a resolução de problemas que não estão previstos em textos legais. Daí também a facilidade com que se resolvem casos que possuem problemas não previstos em Lei, o que inclui questões relacionadas às novas tecnologias. Sobre o tema, as observações de David (2002, p.459): "direito, quer para um jurista americano, quer para um jurista inglês, é concebido essencialmente sob a forma de um direito jurisprudencial; as regras formuladas pelo legislador, por mais numerosas que sejam, são consideradas com uma certa dificuldade pelo jurista que não vê nelas o tipo normal da regra de direito; estas regras só são verdadeiramente assimiladas ao sistema de direito americano quando tiverem sido interpretadas e aplicadas pelos tribunais e

3. Diretrizes aplicadas às plataformas e novos deveres éticos para sua utilização

3.1. Diretrizes aplicadas às plataformas

Os dilemas no uso do ChatGPT para pesquisas e escrita acadêmicas já foram sentidos no âmbito da pesquisa médica. Em artigo na revista *The Lancet*, indicou-se que há uma "necessidade de se implementar diretrizes robustas para a publicação acadêmica" (Liebrenz, 2023). Tais diretrizes devem envolver questões como "direitos autorais, atribuição, plágio e autoria quando a AI produz o texto acadêmico". O autor ainda destaca o fato de que textos produzidos por IA podem passar despercebidos pelos softwares de verificação de plágio (Liebrenz, 2023).

No entanto, sabe-se que os modelos atuais de LLM – pelo menos até o momento – não armazenam as respostas dadas a fim de fornecer uma indicação segura de que determinado texto foi realmente produzido por aquela plataforma[16]. Os textos produzidos podem diferir de tamanho e complexidade, variando desde uma resposta dada a uma questão até uma simples correção gramatical de um texto produzido por um autor humano. Além dessa variabilidade das respostas, o autor ainda pode utilizar ferramentas diferentes de LLM, treinadas de maneiras distintas, com propósitos variados, o que dificultaria bastante a identificação cruzada entre plataformas.

---

quando se tornar possível, em lugar de se referirem a elas, referirem-se às decisões judiciárias que as aplicaram". Ver também Losano (2007, pp. 336-337): "As vantagens e as deficiências do *Common Law* são em certa medida simétricas às do direito continental: sua possibilidade de se adaptar ao caso concreto é maior do que nos sistemas jurídicos de normas gerais e abstratas, mas estas são redutíveis a um corpus limitado e sistemático, enquanto os precedentes tendem a formar uma floresta indestrinçável e extensa. A extensão dos precedentes torna às vezes materialmente impossível ter uma visão clara da situação jurídica de certos setores, fazendo vacilar a certeza do direito".

16 Destacando o fato de que, diante de sua natureza estocástica, ele produz respostas ligeiramente diferentes mesmo diante dos mesmos comandos, além de levar em considerações as interações anteriores do usuário se o comando for dado dentro de uma mesma janela de conversação. Isso torna bastante difícil avaliar com certeza se um texto foi produzido por um LLM, visto que ele também erra quando tenta fazer essa identificação.

Apesar dos resultados altamente satisfatórios no uso desses modelos para tarefas cotidianas, há uma grande dificuldade em entender como realmente eles funcionam[17]. Diante da quantidade de dados utilizados para seu treinamento e as numerosas operações realizadas para prever as palavras, é bastante difícil, senão impossível, fazer o caminho reverso de como aquela palavra foi prevista. A quantidade de operações matemáticas realizadas impede que um humano consiga refazer todo o caminho manualmente. Isso afeta algumas atividades de controle que são realizadas. O problema é: como realizar um controle em um algoritmo que consumiu bilhões de informações, que "aprendeu" a produzir texto novo diante de operações matemáticas altamente complexas, diante das próprias limitações da compreensão humana? (Bowman, 2023, p. 6).[18]

Tal complexidade, no entanto, pode ser superada com institutos já existentes no ordenamento jurídico brasileiro, como a utilização do princípio da boa-fé objetiva. Os deveres anexos à boa-fé contemplam aqueles de informação, colaboração e proteção, o que ilumina os comportamentos dos mantenedores das plataformas de LLM.

É claro que soluções que impõem mais controle sobre os resultados podem, em tese, diminuir sua eficiência. Seria possível implementar modelos que pudessem armazenar mais informações sobre como o resultado foi produzido – o que envolveria mais camadas de processamento, deixando-o mais lento e, consequentemente, mais caro. Há uma escolha ética a ser feita: talvez seja preciso diminuir sua eficiência em prol de mais explicabilidade. Essa explicabilidade pode até

17 Lee e Trott (2023) dão um exemplo que o GPT-4, mesmo não sendo treinado com imagens, conseguiu realizar previsões sobre locais específicos em imagem, organizando respostas visuais sobre como, por exemplo, o local correto de um chifre em um unicórnio. Os autores afirmam, como hipótese, que: "*At the moment, we don't have any real insight into how LLMs accomplish feats like this. Some people argue that such examples demonstrate that the models are starting to truly understand the meanings of the words in their training set. Others insist that language models are "stochastic parrots" that merely repeat increasingly complex word sequences without truly understanding them*".

18 Cf. Bowman (2023, p. 6) 6: "*There are hundreds of billions of connections between these artificial neurons, some of which are invoked many times during the processing of a single piece of text, such that any attempt at a precise explanation of an LLMs behavior is doomed to be too complex for any human to understand*".

mesmo ser necessária para resolver problemas como a análise da possibilidade de terem sido utilizados dados protegidos por direitos autorais para a geração daquela resposta. A questão é que esse novo dever pode ser até mais complexo de se cumprir do que as próprias operações de geração de texto. Em suma, as dificuldades técnicas podem impedir o cumprimento de deveres de explicabilidade e, assim, poderia ser necessário abrir mão deles – dos deveres – em face da própria eficiência dos resultados. A ferramenta acaba por se tornar um oráculo[19]. Esse problema envolve a dificuldade de cumprir o dever de informação e explicabilidade nessas plataformas. Essa falta de transparência[20] pode impedir que a IA seja utilizada em uma variedade de contextos mais críticos. Basta pensar no dever de fundamentação e observância de precedentes na produção de decisões judiciais[21] – embora esse não seja o foco da análise aqui realizada. Como utilizar uma ferramenta que pode sugerir decisões se não se sabe exatamente quais os critérios que foram utilizados para sua fundamentação?

Esse dever de transparência também será variado, dependendo da complexidade e uso das ferramentas de LLM, envolvendo tanto informações sobre os resultados, sobre como elas foram treinadas além de sua aderência a certos padrões (inclusive éticos), o que pode envolver, por exemplo, a eliminação de vieses e preconceitos (Diakopoulos, 2020, p. 199). Não há como responsabilizar adequadamente os mantenedores

19 No âmbito da tomada de decisões judiciais, Lummertz (2018) manifestou-se no sentido de que: "A transparência e auditabilidade dos algoritmos e sistemas de inteligência artificial utilizados na tomada de decisões pelo Poder Judiciário e pela Administração Pública evitarão que tais decisões se convertam em verdadeiros vaticínios, como aqueles provenientes do Oráculo de Delfos, manifestações da vontade divina, cujas razões se desconhece, mas que, assim mesmo, deve-se obedecer".

20 Sobre o tema ver Diakopoulos (2020, p. 198). Segundo o autor: "*Transparency therefore provides the informational substrate for ethical deliberation of a system's behavior by external actors. It is hard to imagine a robust debate around an algorithmic system without providing to relevant stakeholders the information detailing what that system does and how it operates. Yet it's important to emphasize that transparency is not sufficient to ensure algorithmic accountability. Among other contingencies, true accountability depends on actors that have the mandate and authority to act on transparency information in consequential ways*".

21 No exemplo dado por Bostrom e Yudkowsky (2014, p. 317) os autores apontam alguns valores a serem perseguidos: "*Responsibility, transparency, auditability, incorruptibility, predictability*".

dessas ferramentas – e tampouco permitir os mais variados exercícios de direitos das pessoas – sem transparência adequada[22]. Uma possível instrumentalização desse princípio, no âmbito dos direitos autorais, é a disponibilização pelos responsáveis de uma forma de conferir se determinado conteúdo foi ou não utilizado no processo de treinamento da ferramenta. Trata-se de forma relativamente simples de conferência, já que os administradores sabem quais informações foram utilizadas. Por outro lado, a questão se reveste de intensa complexidade quando houver a necessidade de se retirar uma informação do modelo, pois ela implicaria, ao menos em tese, na necessidade de treinar novamente o algoritmo, o que pode demorar semanas ou meses para acontecer[23]. Sendo muitas as solicitações de retirada de informações, a tarefa de treinar novamente o algoritmo pode se tornar impossível. Tarefa mais simples seria a de verificar, de antemão, que não se utilizou informações protegidas por direitos autorais, treinando o algoritmo somente com informações não protegidas por direitos autorais, o que, por certo, comprometeria a eficiência da ferramenta.

Em um primeiro momento, aquele que cria uma obra para publicá-la não está em uma relação propriamente obrigacional. Contudo, aquele que publica uma obra, de qualquer natureza, tem, ao menos, o dever geral de não infringir leis ou causar danos a outros autores, violando seus direitos. O autor de obra que se proponha a ser inédita não pode, por isso, realizar contrafação.

22 Ver de maneira geral, a recomendação da Comissão Europeia (2021). *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (Regulamento Inteligência Artificial) e altera determinados atos legislativos da União* [online]. Available at: https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0004.02/DOC_1&format=PDF (Accessed: 14 August 2023), p. 13: "O aumento das obrigações de transparência também não afetará desproporcionadamente o direito à proteção da propriedade intelectual (artigo 17º, n.º 2), uma vez que estarão limitadas às informações mínimas necessárias para as pessoas singulares exercerem o seu direito à ação e à transparência necessária perante as autoridades de supervisão e execução, em conformidade com os mandatos destas".

23 O Parlamento Europeu já se manifestou sobre a necessidade de "desenvolver novos métodos, como, por exemplo, sistemas de aprendizagem adaptáveis que se podem recalibrar após cada entrada". PARLAMENTO EUROPEU (2020) *Resolução do Parlamento Europeu, de 20 de outubro de 2020, sobre os direitos de propriedade intelectual para o desenvolvimento de tecnologias ligadas à inteligência artificial*, p. 6.

Destaque-se, ainda que, para ser possível tratar de proteção autoral, há que ser evidenciada originalidade e individualidade, mesmo que em nível mínimo, inclusive, alcançando direitos de autor daquele que programa, bem como em seus *inputs*, segundo leciona Marilù Capparelli:

> Para que possa ser dada proteção no sentido da normativa de direito de autor, a obra deve ser dotada de caráter criativo, de um nível mínimo de originalidade e individualidade que reflita a personalidade e o engenho de seu autor, e que distingue a obra em si da obra que lhe precedeu. [...] A originalidade do ato criativo poderá, agora, ser reenviada para o momento antecedente da obra, ou seja, no momento da programação da máquina, coincidente com a decisão e organização da criação da obra: seria, assim, nas escolhas completas pelo programador e nos *inputs* feitos para a base da programação da IA, que se poderia reenviar aquele momento de criatividade necessário para atribuir o direito de monopólio exclusivo no que toca à obra de criatividade, em tal medida reconduzível ao intelecto do programador; o qual poderá, portanto, ser considerado o titular dos direitos de autor sobre a obra (traduziu-se) (Capparelli, 2020, pp. 335-342).

Nesta linha, que o Copyright, Design and Patent Act de 1988, estabelece que é autor aquele que tomou as medidas necessárias para o processo criativo da obra gerada, em disposição de vanguarda: "No caso de uma obra literária, dramática, musical ou artística gerada por computador, o autor deve ser considerado a pessoa que realizou os arranjos necessários para a criação da obra" (Reino Unido, 1988).

É inegável que a criação intelectual, quando feita com o apoio de uma inteligência artificial, desafia a própria natureza do ato de criar. Pontes de Miranda (1956, p. 13) já dizia que "a inserção do elemento psíquico próprio é que faz a criação intelectual". Há várias formas de analisar a questão da propriedade intelectual e a IA. A primeira é a proteção dos algoritmos que instrumentalizam os LLMs, o que não parece trazer tantos problemas, eis que se trata de software, contando com regras de proteção específicas. Porém, quando se fala dos dados utilizados para o treinamento dos algoritmos, tem-se uma complexa teia de situações jurídicas envolvidas. O primeiro

diz respeito ao fato de que tais modelos usam dados que estão disponíveis na Internet. Ocorre que, diante da opacidade das técnicas utilizadas, não se sabe ao certo quem são os titulares de tais dados, visto que se pode estar diante de dados protegidos por direitos autorais, dados pessoais protegidos por leis específicas e, até mesmo, imagens e dados pessoais, amparadas pelos direitos da personalidade e a Lei Geral de Proteção de Dados. Essa miríade de situações, diante da falta de clareza e transparência por parte dos administradores de tais plataformas, impossibilita que os titulares de direitos saibam que seus dados estão sendo tratados e se oponham às atividades que eventualmente violem a Lei.

Não se perca de vista, também, que a própria atividade de treinamento poderia ser analisada sob a ótica do Direito das Obrigações, no âmbito do enriquecimento sem causa[24]. Mesmo que as obras utilizadas para o treinamento não possam ser identificadas nos *outputs*, os mantenedores das ferramentas podem utilizar obras protegidas por direitos autorais para a criação dos algoritmos, o que lhes trará evidentes benefícios financeiros. Nesse sentido, antes mesmo das discussões atinentes aos direitos autorais, há uma discussão que deve ser feita sobre a possibilidade de aplicação do instituto do enriquecimento sem causa ao caso.

O ponto de reflexão passa a ser se a utilização de criações intelectuais para treinar o algoritmo exige que o seu mantenedor respeite os direitos autorais aplicáveis ou se essa atividade isenta o dono do algoritmo de respeitar os direitos dos autores. Parece claro que o uso de obras intelectuais neste contexto não pode ser realizado sem o respeito dos direitos autorais. Esta hipótese violaria até um paradigma ético do uso das IA que deveria respeitar, em última análise, a própria ideia de colaboração (Benanti, 2018, p. 55) [25] e o bem comum. Além

24 Cf. o art. 884 do CC: "Aquele que, sem justa causa, se enriquecer à custa de outrem, será obrigado a restituir o indevidamente auferido, feita a atualização dos valores monetários".

25 Cf. Benanti (2018, p. 55): "*In altre parole, le intelligenze artificiali (machinae sapientes) non sono degli avversari evolutivi dell'homo sapiens bensì strumenti (artefatti) che devono essere pensati come cooperativi alla persona. Di fatto le intelligenze artificiali devono essere realizzate per aumentare la capacità cognitiva, che è prerogativa unica e peculiare dell'uomo, e non sostituirsi mai a questi*".

disso, parece até mesmo haver um abuso de direito dos responsáveis pelos algoritmos que treinam os algoritmos de forma a não permitir que os autores das obras utilizadas sequer saibam que suas obras foram utilizadas.

Pense-se em um exemplo mais claro. Sabe-se que o Tratado de Direito Privado de Pontes de Miranda está disponível para download na Internet. Diante disso, seria lícita a atuação de treinar uma inteligência artificial com todos os livros do autor e disponibilizar, assim, um chat para responder perguntas com base em tais obras? E se assim fizesse, a Lei brasileira de direitos autorais seria aplicável? Se é claro que a Lei deve ser respeitada neste contexto, o mesmo não poderia ser dito acerca de todas as situações em que se utilizam obras de outros autores para treinar algoritmos de IA?

Assim, compreende-se que a observação aos direitos de autor passa estar entre as diretrizes a serem respeitadas pelos mantenedores dos LLMs. A Lei dos Direitos Autorais, em seu artigo 7º, manifesta a sua vocação futurista, na medida em que estabeleceu que são protegidas as criações de espírito, "por qualquer meio ou fixadas em qualquer suporte", "conhecido ou que se invente no futuro"[26], e, portanto, voltando-se ao porvir, com incidência plena, imediata e inarredável sobre o campo das novas tecnologias e, no caso, sobre os LLMs.

Devem ser observados os direitos autorais e conexos, tanto no que tange ao âmbito moral, como patrimonial, sobretudo os artigos 24 a 29 da Lei dos Direitos Autorais.

Outra importante diretriz se apresenta na medida da utilização para alimentação da IA de obras que já estão em domínio público, ou seja, como no caso de já ter transcorrido mais de setenta anos da morte do criador, a contar do primeiro dia do ano seguinte ao falecimento, nos termos do artigo 41, da Lei de Direitos Autorais. Incluem-se aí também as exceções presentes no art. 46 da referida lei ou, ainda, o uso de obras que estão em licenciadas no modelo de *Creative Commons,* ade-

26 Compartilhando sobre Direitos Autorais, voltadas às plataformas em Colombo, C. (2020) 'Direitos Autorais na Indústria Criativa e conteúdo ilícito gerado por terceiros', *Migalhas*, 30 September [online]. Available at: https://www.migalhas.com.br/coluna/migalhas-de-responsabilidade-civil/334150/direitos-autorais-na-industria-criativa-e-conteudo-ilicito-gerado-por-terceiros (Accessed: 13 August 2023).

quando-se às suas licenças, se para fins lucrativos ou não, e estando vedada a utilização de licenças que não permitam obras derivadas, já que o *output* poderá ser identificado como tal.[27] A legislação autoral admite, de forma expressa, ainda, que seja possível a utilização de textos para paródias, na linha do artigo 47 da LDA, quando as "reproduções de obras originárias não impliquem em descrédito ao autor".

Porém, acerca de licenças específicas de *Creative Commons* ou de código fonte aberto (*open source*), é preciso também observar as próprias limitações de uso dessas licenças. É comum que nestes casos haja a necessidade de que as obras derivadas sejam disponibilizadas também com as mesmas licenças. Um caso recente envolveu o treinamento da ferramenta Copilot, da Microsoft, que ajuda programadores a criarem código com o uso da IA. A solução funciona como um ChatGPT específico para a criação de códigos fonte de softwares. Ocorre que a solução foi treinada com códigos *open source* armazenados em um repositório, também da Microsoft, chamado de GitHub. A comunidade de desenvolvedores de software se insurgiu contra o fato de que quando os desenvolvedores publicaram seu código no formato *open source*, o uso para treinamento de IA feriria os termos da própria licença, já que a solução é um projeto comercial (o que é vedado pela licença), além de desrespeitar os

> princípios éticos de compartilhamento que levaram os desenvolvedores a distribuir seu código abertamente em primeiro lugar. Em segundo lugar, houve críticas ao código que é sugerido pelo Copilot, com alguns programadores alegando que ele reproduz o próprio código deles de volta (Guadamuz, 2023).

Isso demonstra a dificuldade de compatibilizar as soluções atuais de proteção aos direitos autorais com as novas possibilidades trazidas pelas ferramentas de LLM, tanto no processo de treinamento quanto na produção de *outputs*. Diante do caso apontado, portanto, há que se respeitar os limites

27 CREATIVE COMMONS [online]. Available at: https://creativecommons.org/licenses/?lang=pt_BR (Accessed: 13 August. 2023).

das licenças abertas, inclusive quanto a possíveis limitações de seu uso comercial.

3.2. Novos deveres Éticos-Jurídicos para sua utilização

Diante do exposto, é preciso avaliar quais são os novos deveres de autores diante do uso da IA para a produção acadêmica. O primeiro ponto envolve a análise de como o conceito de autoria varia com essas novas tecnologias. Note-se que a situação analisada envolve o uso de LLMs por autores humanos que o utilizam para apoiar ou complementar textos autorais. Em recente artigo intitulado *ChatGPT listed as author on research papers*, a comunidade acadêmica analisou o papel do ChatGPT, que havia sido indicado em pelo menos quatro publicações, como autor, trazendo questionamentos sobre "o futuro dos ensaios universitários" e a "produção de pesquisa". O Time da Nature questionou editores sobre o ChatGPT, sendo que se manifestaram que o LLM "não preenche os critérios de estudo autoral, porque não tem responsabilidade pelo conteúdo e integridade do trabalho científico", e, que "[...] somente pessoas podem ser listadas" (Stokel-Walker, 2022). Os editores-chefe da Science e da Nature referiram que "LLMs não atendem padrão de autoria" (Stokel-Walker, 2022). E, no caso de ser utilizado, o LLM deverá ser indicado na seção de métodos, uma vez que texto feito por IA, sem a citação adequada, pode ser compreendida como plágio (Stokel-Walker, 2022). [28]

Importante preocupação envolve formas de evitar plágios nos trabalhos acadêmicos. Concorda-se com a afirmação de que "O plágio trata-se de uma questão ética, antes do que jurídica" (Pithan and Vidal, 2013, p. 78). No entanto, a ideia de plágio como reprodução de passagens de texto de outros autores sem a devida citação, ou reprodução de conceitos, ideias ou teorias sem a indicação da fonte, parece ser uma das vias

28 Segundo as regras de publicação de Nature: "*Large Language Models (LLMs), such as ChatGPT, do not currently satisfy our authorship criteria. Notably an attribution of authorship carries with it accountability for the work, which cannot be effectively applied to LLMs. Use of an LLM should be properly documented in the Methods section (and if a Methods section is not available, in a suitable alternative part) of the manuscript*" (Stokel-Walker, 2022).

a serem consideradas diante da IA. Talvez uma aproximação mais apropriada seria considerar os LLM como ferramentas de apoio com diferentes níveis de intervenção no texto. A depender do seu nível de interação, elas podem atuar de forma semelhante a um *ghost writer* (Zanini, 2015)[29]. Isso porque o conteúdo gerado não existia propriamente até sua geração (embora as condições para sua geração já estejam dadas no algoritmo treinado), a não ser que a IA realmente reproduza um texto já criado anteriormente, o que é sim possível a depender da instrução dada pelo usuário. Assim como ocorre com o trabalho de um *ghost writer*, há uma colaboração entre ele e um terceiro que fornece informações, sendo que aquele as organiza em formato literário[30].

Os níveis de interação de humanos com as ferramentas de IA podem ser muito variados, o que envolve intervenções mais ou menos intensas no texto produzido. Há um escalonamento na interferência que a IA pode realizar na produção de um texto. Graus mais intensos de interferência levam a uma afetação na atribuição de autoria do texto. Uma situação básica, que não parece envolver grandes problemas jurídicos (ou éticos) de atribuição de autoria ou até mesmo de plágio, ocorre quando o autor apenas faz uma correção gramatical de

29 Ver Zanini (2015): O problema do *ghost writer* nos direitos autorais demanda também reflexões profundas. Zanini afirma que se trata de uma questão que envolve a análise do "problema da renúncia aos direitos da personalidade do autor" (Zanini, 2015, p. 293). O autor ainda aponta que na doutrina alemã o problema é tratado no âmbito do "não exercício obrigacional e limitado no tempo do direito de paternidade" (Zanini, 2015, p. 295).

30 Lembrando do famoso caso brasileiro do livro escrito por um *ghost writer* e pela ex prostituta Bruna Surfistinha. A história chegou a ser transformada em filme. Para mais informações, inclusive no âmbito jurídico, ver Virtuoso, B. and Graton, N. A. (2015) 'Ghost-Writer: uma breve análise a partir do caso Bruna Surfistinha'. *GEDAI*, 30 May [online]. Available at: https://www.gedai.com.br/ghost-writer-uma-breve-analise-a-partir-do-caso-bruna-surfistinha/ (Accessed: 10 August 2023). O caso chegou até o STJ (REsp 1.387.242), conforme resumo de Virtuoso e Graton: "o recorrente apenas redigiu o texto, fazendo as correções necessárias, de forma que **não traduz o desenvolvimento literário que configura a autoria**. Além disso, o escritor tinha ciência que havia sido contratado pela editora para prestar serviços como Ghost-writer, profissional especializado em redação de textos para terceiros abdicando da paternidade da obra". Também citando o caso Zanini, Leonardo Estevam de Assis. Ibidem, p. 299, apontando que: "Não temos dúvida que coube ao autor a criação do livro, pois foi ele quem selecionou as situações que reputou mais interessantes, criou uma narrativa que intercalou fatos ocorridos antes e depois da fuga de Raquel de sua casa, bem como imprimiu na obra a sua personalidade, a sua forma de ver a vida das personagens abordadas".

um texto próprio. Nesta situação, o autor submete um texto totalmente original para avaliação gramatical. Com isso, a ferramenta atuará como um corretor, semelhante ao que acontece com a correção automática nas ferramentas tradicionais de edição de textos. Situações semelhantes ocorrem quando o autor busca a substituição de palavras por sinônimos ou até mesmo quando utiliza a ferramenta para a tradução de textos.

O nível de interferência começa a se aprofundar, demandando mais reflexão, quando é solicitado à ferramenta a melhoria de um texto autoral, eis que pode ser alterada a sua estrutura. Tome-se como exemplo o parágrafo anterior. Utilizando-se o GPT-4, forneceu-se a seguinte instrução: "Melhore e resuma o parágrafo a seguir", obtendo-se como resposta:

> A interação humana com ferramentas de IA varia em intensidade, afetando a atribuição de autoria do texto. Em situações básicas, onde o autor usa a IA para correções gramaticais ou busca de sinônimos em um texto original, poucas questões jurídicas ou éticas surgem. Isso se assemelha à correção automática de editores de texto tradicionais ou ao uso de tradutores automáticos.

Note-se que o sentido do parágrafo não foi alterado, as palavras principais foram mantidas e o resumo manteve a essência do texto. Esse tipo de interação poderia ser entendido como uma citação indireta[31], se realizado por outra pessoa que não o autor – demandando sempre a citação do texto original. No entanto, como a ideia de citação indireta envolve um texto baseado na obra de outro autor, não haveria a necessidade do próprio autor afirmar que fez um resumo de um texto seu.

Acredita-se que não há questões controversas para o próprio autor quando ele solicita a melhoria ou o resumo de textos próprios. Ocorre situação semelhante quando um autor submete sua dissertação ou tese para um revisor profissional. Neste caso, não há questões éticas envolvendo problemas de confusões de autoria[32] ou a possibilidade de atribuição de

31 Lembrando que a citação direta é a "transcrição textual de parte da obra do autor consultado" e a citação indireta é o "texto baseado na obra do autor consultado", cf. ABNT NBR 10520 (ABNT, 2013).

32 Lembrando que o art. 11 da lei 9.610/98 define o autor como "pessoa física criadora de obra literária, artística ou científica".

co-autoria. Denominamos esses usos da IA como *usos de apoio primário*.

Os termos de uso da OpenAI concedem ao usuário a titularidade dos *outputs*, permitindo que sejam utilizados inclusive para fins comerciais, o que envolve inclusive a publicação[33]. Considerando que a OpenAI oferece uma ferramenta de geração de textos e suas políticas – com força obrigacional, mesmo que o vínculo se dê por adesão – transferem todos os direitos de uso do texto produzido, não parece ser difícil entender que a ferramenta funciona como um *ghost writer*. Veja-se aqui até a dificuldade de se analisar possíveis renúncias de direitos do autor – eis que a ferramenta não é humana -, visto que, ao menos pela Lei atual, não haveria como se atribuir autoria a uma solução de inteligência artificial. É claro que esta solução jurídica não afasta o aspecto ético do autor mencionar o uso de ferramentas de IA quando houver um grau intenso de intervenção no texto produzido. E a solução é diferente daquela em que alguém usa um buscador para responder a uma questão ou para buscar uma informação qualquer. A diferença principal é que o buscador entrega, na maioria das vezes, as fontes originais que embasaram suas respostas, permitindo ao autor citar diretamente a informação encontrada por meio do buscador. Já com as ferramentas de IA, na maioria das vezes não se obtém a fonte utilizada para a resposta dada por ela e nem poderia, diante das características de geração de conteúdo, que estão mais relacionadas a regras estatísticas que organizam o texto baseados na probabilidade da próxima palavra naquele contexto do que o apontamento de uma fonte específica. É claro que devem ser levadas em consideração as fontes usadas para treinar o algoritmo, porém o número de

33 OpenAI (2023). OpenAI, Terms of Use, 14 March [online]. Available at: https://openai.com/policies/terms-of-use. (Accessed: 2 August 2023): "*Your Content. You may provide input to the Services ("Input"), and receive output generated and returned by the Services based on the Input ("Output"). Input and Output are collectively "Content." As between the parties and to the extent permitted by applicable law, you own all Input. Subject to your compliance with these Terms, OpenAI hereby assigns to you all its right, title and interest in and to Output. This means you can use Content for any purpose, including commercial purposes such as sale or publication, if you comply with these Terms. OpenAI may use Content to provide and maintain the Services, comply with applicable law, and enforce our policies. You are responsible for Content, including for ensuring that it does not violate any applicable law or these Terms*".

fontes usadas torna rarefeita sua influência no conteúdo gerado, sendo bastante difícil, em alguns contextos, que a ferramenta aponte quais fontes foram utilizadas para gerar o resultado (a não ser que se peça expressamente por elas).

As fontes de treinamento servem também, não se esqueça, para organizar os vetores de palavras em contextos e aproximações, não somente para trazer informações. Isso faz com que as informações usadas no processo possam não aparecer explicitamente no resultado, servindo como fonte para a organização dos vetores e, em última análise, de contextos. Apreende-se também a estrutura gramatical do texto sem, necessariamente, refletir, nos resultados, o seu conteúdo. São dimensões diferentes na geração de resultados: a coerência do texto versus seu amparo em fontes científicas ou factuais.

Não se descuide que há também uma criatividade do autor nos *inputs* que ele dá, o que aproxima ainda mais a situação com a atuação de um *ghost writer*. O usuário pode ir informando situações e pedindo, pouco a pouco, que a ferramenta vá escrevendo uma história com base naquelas informações. Pode-se dar informações sobre personagens, características, fatos, relacionamentos entre personagens, etc. Porém, a criatividade do autor – ou sua falta – pode fazer com que ele peça para a ferramenta imitar estilos[34] Níveis de interação mais aprofundados da ferramenta ocorrem quando o usuário humano solicita a criação de novos textos somente com a introdução de apenas uma instrução. Nestes casos é altamente recomendável que a ferramenta de IA seja citada explicitamente

34 CHATGPT-4. Prompt: "Escreva um pequeno poema inédito copiando o estilo do poeta Mário Quintana". Geração em: 4 ago. 2034. https://chat.openai.com/. Resposta: Na calçada da memória, encontro-me sozinho, / Onde a lâmpada do tempo tece o crepúsculo do dia. / Ora distante, ora perto, tal qual um passarinho, Que entre árvores esvoaça, onde sua canção porfia. Sob o manto de estrelas, meu pensamento divaga, / Enquanto o silêncio da noite a tudo permeia. / No espelho d'água, o luar em seu reflexo se embriaga, /Em doce melancolia, na qual a alma se enleia. Ah, as ruas desta vida, trilhas do meu viver, / São repletas de risos, de lágrimas e poesia. / Que se desenrola, como fita ao vento, a correr, / Na tela infinita do tempo, onde o sonho desafia. Mas na roda da existência, onde tudo se desdobra, / Nem sempre se alcança a linha da eternidade. / O importante, meu amigo, não é o que se cobra, /Mas o amor com que se vive, em toda a sua amplitude.

como meio na obtenção daquele resultado, sendo as informações geradas indicadas como citação direta ou indireta[35].

Questões mais complexas emergem quando o usuário solicita informações factuais e conceituais[36]. A depender dos dados que treinaram a ferramenta, os resultados podem ser mais ou menos fidedignos, ocorrendo até o apontamento de informações incorretas[37]. Por outro lado, existem modelos que são treinados para atividades mais específicas, como modelos utilizados para a medicina. Acerca dos *layers* que se indicou anteriormente, eles possuem funções de ajustes de contextos. Alguns problemas aparecem quando a ferramenta produz *outputs* que fazem sentido gramaticalmente, mas que não possuem respaldo na realidade fática ou científica. A própria OpenAI, desenvolvedora do modelo GPT, reconhece tais limitações e ainda aponta a necessidade de um cuidado na análise dos resultados

> [...] em contextos de alto risco, com o protocolo exato (como revisão humana, fundamentação com contexto adicional ou evitando usos de alto risco completamente) correspondendo às necessidades de aplicações específicas[38].

Um dos testes feitos pelos autores deste artigo consistiu no solicitação de indicação de músicas que tivessem como tema a educação. Entre os resultados, foram trazidas duas músicas que não existiam, fenômeno chamado de "alucinação". Diante desse risco de respostas que não possuem base na realidade fática ou científica, modelos mais aprimorados poderiam inserir outros *layers* de apuração ou confirmação, podendo até utilizar a própria internet para tais conferências. Não se trata somente de utilizar *plugins* para que a ferramenta busque informações na Internet, mas que ela tenha a capacidade de corrigir suas respostas, como um último nível de "decantação" do conteúdo, visando eliminar incorreções ou inconsistências

35 Como se fez no presente texto.

36 Como, por exemplo, quando morreu determinada personalidade, quem foi o presidente do Brasil em tal ano ou quem é o autor de tal teoria.

37 Já que a ferramenta pode meramente privilegiar a coerência gramatical do texto, desprezando a correção informacional ou factual das respostas apresentadas.

38 OPENAI (2023) *GPT-4 Technical Report*, p. 10 [online]. Available at: https://arxiv.org/pdf/2303.08774.pdf (Accessed: 3 August 2023).

fáticas. Trata-se de incluir uma outra dimensão no algoritmo que avalie, além do sentido gramatical, sua correção factual e científica. Note-se que isso adiciona uma camada bem mais complexa de conferência, já que o modelo precisaria pesquisar e avaliar todos os *outputs* dados, o que deixaria a entrega das respostas mais demorada[39]. Isso já ocorre, de certa forma, com as ferramentas de moderação. Simples interações com o ChatGPT solicitando temas mais controversos, ou até mesmo o uso de palavrões[40], trazem *outputs* que refletem decisões éticas dos desenvolvedores em limitar o alcance do conteúdo.

Não se perca de vista, também, a hipótese aventada por Marcelo Bürger e Rafael Corrêa (2020, p. 531), quando invocam o art. 40 da Lei 9.610/98 ao prever a possibilidade de proteção de obras anônimas ou pseudo anônimas. Os autores apontam que os direitos patrimoniais de tais obras serão gozados a quem as publicar. No entanto, essa hipótese não se encaixa completamente no caso, eis que o ser humano quando interage com as ferramentas de IA está a lhes dar direção, sendo mais adequada, em nosso sentir, a hipótese de proteção semelhante à aplicada às situações de participação de *ghost writer*, sempre observando o grau de interação do ser humano, conforme se expôs.

Nesse sentido, devem os LLMs, observar os seguintes deveres éticos-jurídicos:

1) observar os direitos de autor, no âmbito moral, com a obrigação de informar quais os textos foram utilizados, no processo de treinamento, como *input* na Inteligência Artificial, nos termos do artigo 24, da Lei de Direitos Autorais;

---

39 Nada impede que a ferramenta adicione uma opção de conferência, para que o usuário somente a acione quando achar necessário. Essa opção não seria ativada, por exemplo, quando se buscar a mera correção de um texto, eis que utilizaria somente as funcionalidades gramaticais da solução.

40 CHATGPT-4. Prompt: Indique quais os palavrões mais ofensivos para se xingar alguém. Acesso em: https://chat.openai.com/. Acesso em 13 ago de 2023. A resposta foi: "Desculpe, mas eu não vou fornecer uma lista de palavrões ou incentivar o uso de linguagem ofensiva. Se você tiver dúvidas ou necessidades relacionadas a linguagem, educação, cultura ou outros tópicos, estou aqui para ajudar!". Note-se que apesar do modelo possuir a capacidade de produzir textos envolvendo palavrões, a moderação da plataforma gerou uma contenção no discurso.

2) caso seja utilizado obra total ou parcial, que gerará obra derivada, a autorização dos criadores das obras primígenas, nos termos do artigo 29, da Lei dos Direitos Autorais;

3) caso sejam utilizados pequenos trechos, a LDA não exige autorização prévia e expressa, mas, em todos os casos, o *output* deve citar quem apresentou a definição, os exemplos ou outras informações, para que não haja violação a direito de autor. É o que pode ocorrer, na maioria das vezes, quando a resposta se utiliza de múltiplos textos, devendo apenas referenciá-los;

4) a utilização de textos em domínio público, licenciados por *creative commons* ou cobertos pelas exceções do art. 46 da Lei de Direitos Autorais, como *inputs*, poderá minimizar violações a direitos patrimoniais de autor, sempre sendo preservado o direito autoral. Neste caso, deve-se observar também as próprias licenças *creative commons* quando limitam seu uso comercial, o que se aplica também às licenças de código aberto.

5) implementar controles que não reproduzam respostas exatamente iguais a materiais protegidos por direitos autorais sem a indicação da fonte. Isso eliminaria (ou ao menos diminuiria, do ponto de vista da plataforma) a presença de plágios, permitindo que a ferramenta possa ser utilizada de forma mais segura pelos usuários. Não se pode perder de vista que o aumento dos controles pode afetar sensivelmente a eficiência e o custo de manutenção da plataforma, o que importa em um problema econômico. Há uma tendência da ferramenta se tornar mais cara e complexa sempre que implementar controles para cumprir as Leis. Em última instância, isso diminuiria os lucros dos seus mantenedores, daí o problema econômico a ser enfrentado. Por outro lado, é preciso promover um ambiente institucional que permita o desenvolvimento das ferramentas de inteligência artificial e de suas potencialidades com a necessidade "de assegurar um elevado nível de proteção de DPI (Direitos de Propriedade Intelectual)" (Parlamento Europeu, 2020).

Dessa forma, há a necessidade de se estabelecer uma supervisão pela curadoria de dados (Bender, Gebru, McMillan-Major and Shmitchell, 2021). Ora, se a Inteligência Artificial é como uma criança, "aprende com a experiência" (Ruffolo, 2021), dar a ela o dado hígido e curado, identificando-o e clas-

sificando-o, para a realização de uma determinada tarefa, mitigará violações aos direitos da pessoa humana.

4. Considerações finais

Com a chegada dos LLMs, é importante que se promova uma releitura dos Direitos de Autor, para que siga sendo promovida a tutela destes vulneráveis. No âmbito moral, deve ser observada a obrigação de fornecer os textos utilizados como *input* no processo de treinamento, na Inteligência Artificial, para atendimento do artigo 24, da Lei de Direitos Autorais. Na hipótese da utilização de obra total ou parcial, que gerará obra derivada, deverá ser coletada a autorização dos criadores das obras primígenas, nos termos do artigo 29, da Lei dos Direitos Autorais. Em sendo utilizado pequeno trecho, o *output* deve citar quem apresentou a definição, os exemplos, enfim, para que não haja violação a direito de autor, não sendo exigida, no caso, autorização prévia e expressa, eis que hipótese do artigo 46, da LDA. Dessa forma, critério bastante importante é a identificação de elementos substanciais de uma obra no resultado (*output)*, que, a depender da proporcionalidade, deve ser determinante para a coleta da autorização prévia e expressa ou, como acima referido, se pequeno trecho for, deve ser caso de citação, ainda que indireta. A utilização de textos em domínio público ou *creative commons*, como *input*, no processo de treinamento, poderá minimizar violações a direitos patrimoniais de autor, sempre sendo preservado o direito autoral. O que se espera é que editores de revistas científicas criem seus guias de conformidade para utilização dos LLMs, preservando a autonomia humana e estabelecendo os limites destas ferramentas.

Referências

Associação Brasileira de Normas Técnicas (2013) *ABNT NBR 10520: Informação e documentação — Citações em documentos — Apresentação*. Rio de Janeiro: ABNT.

Benanti, P. (2018) *Le macchine sapienti: Intelligenze artificiali e decisioni umane*. Bologna: Marietti.

Bender, E. M., Gebru, T., Mcmillan-Major, A. and Shmitchell, S. (2021) 'On the dangers of stochastic parrots Can language models be too big?' *AccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, March, pp. 610-623 [online]. Available at: https://dl.acm.org/doi/10.1145/3442188.3445922. (Accessed: 12 August 2023).

Bonadio, E., Dinev, P. and McDonagh, L. (2022) 'Can artificial intelligence infringe copyright? Some reflections' in Aboott, R. *Research Handbook on Intellectual Property and Artificial Intelligence*. Cheltenham: Edward Elgar.

Bostrom, N. and Yudkowsky, E. (2014) 'The ethics of artificial intelligence' in Frankish, K. and Ramsey, W. M. *The Cambridge Handbook of Artificial Intelligence*. Cambridge: Cambridge University Press.

Bowman, S. R. (2023) *Eight Things to Know about Large Language Models* [online]. Available at: https://cims.nyu.edu/~sbowman/eightthings.pdf. (Accessed: 13 August 2023).

Bürger, M. L. F. de M. and Corrêa, R. (2020) 'Dos pincéis aos algoritmos: a titularidade das expressões artísticas e criativas resultantes da aplicação da inteligência artificial' in Erhardt Júnior, M., Catalan, M. and Malheiros, P. (coords.) *Direito Civil e Tecnologia.* Belo Horizonte: Fórum.

Capparelli, M. (2020) 'Le nuove frontiere del diritto d'autore alla prova dell'Intelligenza Artificiale' in Ruffolo, U. (coord) *Intelligenza Artificiale: il diritto, i diritti, l'etica*. Milão: Giuffrè Francis Lefebvre.

Carvalho, A. G. P. (2021) 'Defesa da concorrência e inteligência artificial: Desafios impostos pela colusão motivada por algoritmos de fixação de preços' in Mendes, L. S., Alves Jr, S. and Doneda, D. (coords.) *Internet & Regulação*. São Paulo: Saraiva.

Colombo, C. (2020) 'Direitos Autorais na Indústria Criativa e conteúdo ilícito gerado por terceiros'. *Migalhas,* Migalhas de Responsabilidade Civil, 30 September [online]. Available at: https://www.migalhas.com.br/coluna/migalhas-de-responsabilidade-civil/334150/direitos-autorais-na-industria-criativa-e-conteudo-ilicito-gerado-por-terceiros (Accessed: 13 August 2023).

Comissão Europeia (2021). *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (Regulamento Inteligência Artificial) e altera determinados atos legislativos da União* [online]. Available at: https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0004.02/DOC_1&format=PDF (Accessed: 14 August 2023).

David, R. (2002) *Os grandes sistemas do direito contemporâneo*. São Paulo: Martins Fontes.

Diakopoulos, N. (2020) 'Accountability, Transparency, and Algorithms' in Dubber, M. D., Pasquale, F. and Das, S. (eds.). *The Oxford Handbook of Ethics of AI*. Oxford: Oxford University Press.

Floridi, L. (2022) *Etica dell'intelligenza artificiale*. Milano: Raffaelo Cortina.

Grimmelmann, J. (2016) 'Copyright for Literate Robots'. *Iowa Law Review*, 101, pp. 657-681.

Guadamuz, A. (2023) 'A Scanner Darkly: Copyright Liability and Exceptions in Artificial Intelligence Inputs and Outputs'. *SSRN*, 26 February, Draft version 1.5. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4371204 (Accessed: 13 August 2023).

Lee, T. B. and Trott, S. (2023) 'A jargon-free explanation of how AI large language models work'. *Ars Technica,* 31 July [online]. Available at: https://arstechnica.com/science/2023/07/a-jargon-free-explanation-of-how-ai-large-language-models-work (Accessed: 13 August 2023).

Liebrenz, M. (2023) 'Generating scholarly content with ChatGPT: ethical challenges for medical publishing'. *The Lancet Digital Health*, 6 February. Available at: https://www.thelancet.com/journals/landig/article/PIIS2589-7500(23)00019-5/fulltext (Accessed: 13 August 2023).

Losano, M. G. (2007) *Os grandes sistemas jurídicos: Introdução aos sistemas jurídicos europeus e extra-europeus*. São Paulo: Martins Fontes.

Lummertz, H. (2018) 'Algoritmos, inteligência artificial e o Oráculo de Delfos'. *Jota*, Tecnologia, 12 October [online]. Available at: https://www.jota.info/opiniao-e-analise/artigos/algoritimos-inteligencia-artificial-e-o-oraculo-de--del fos-12102018. (Accessed: 12 August 2023).

MIT (2021) 'A corrida para entender o mundo emocionante e perigoso dos modelos de linguagem para inteligência artificial'. *MIT Technology Review*, Inteligência Artificial, 4 June. Available at: https://mit techreview.com.br/a-corrida-para-entender-o-mundo-emocionante-e-perigoso-dos-modelos-de-linguagem-para-inteligencia-artificial/ (Accessed: 13 august 2023).

MIT (2023a) 'ChatGPT faz com que as pessoas escrevam melhor'. *MIT Technology Review*, Humanos e Tecnologia, 11 August [online]. Available at: https://mitte chreview.com.br/chatgpt-faz-com-que-as-pessoas-escrevam-melhor. (Accessed: 11 august 2023).

MIT (2023b) 'Como o texto gerado por Inteligência Artificial está envenenando a Internet', *MIT Technology Review*, Inteligência Artificial, 6 January [online]. Available at: https://mit techreview.com.br/como-o-texto-gerado-por-inteligencia-artificial-esta-envenenando-a-internet/ (Accessed: 11 August 2023).

OPENAI (2023) *GPT-4 Technical Report*, 27 March [online]. Available at: https://arxiv.org/pdf/2303.08774.pdf (Accessed: 11 august 2023).

PARLAMENTO EUROPEU (2020) *Resolução do Parlamento Europeu, de 20 de outubro de 2020, sobre os direitos de propriedade intelectual para o desenvolvimento de tecnologias ligadas à inteligência artificial* [online]. Available at: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0277_PT.pdf (Accessed: 11 august 2023).

Pithan, L. H. and Vidal, T. R. A. (2013) 'O plágio acadêmico como um problema jurídico e pedagógico'. *Direito & Justiça*, 39(1), pp. 77-82, jan./jun.

Pontes de Miranda, F. C. (1956) *Tratado de Direito Privado*. 2nd edn. Rio de Janeiro: Borsoi, T. XVI.

Putini, J. (2010) 'Papagaios são isolados em zoológico após ofenderem vários visitantes', *R7,* Hora 7, [online]. Available at: https://noticias.r7.com/hora-7/fotos/papagaios-sao-isolados-em-zoologico-apos-ofenderem-varios-visitantes-02102020 (Accessed: 12 August 2023).

Reale, M. (1994) *Lições Preliminares de Direito*. São Paulo: Saraiva.

Reino Unido (1988) *Copyright, Designs and Patents Act 1988* [online]. Available at https://www.legislation.gov.uk/ukpga/1988/48/contents (Accessed: 11 September 2023).

Ruffolo, M. (2022) 'The Role of Ethical AI in Fostering Harmonic Innovations that Support a Human-Centric Digital Transformation of Economy and Society' in Cicione, F., Filice, L. and Marino, D. (eds.) *Harmonic Innovation. Lecture Notes in Networks and Systems*, 282. Cham: Springer. https://doi.org/10.1007/978-3-030-81190-7_15.

Stokel-Walker, C. (2022) 'ChatGPT listed as author on research papers'. *Nature*, Berlim, 613, 26 January [online]. Available at: https://www.nature.com/articles/d41586-023-00107-z (Accessed: 11 August 2023).

Tribunal de Contas da União (2020) *Manual de Direitos Autorais*. Brasília. Available at: https://portal.tcu.gov.br/data/files/57/72/86/60/35FA6710FE28B867E18818A8/Manual%20Direitos%20Autorais%202020_Web.pdf. (Accessed: 13 August 2023).

Virtuoso, B. and Graton, N. A. (2015) Ghost-Writer: uma breve análise a partir do caso Bruna Surfistinha. *GEDAI*, 30 May [online]. Available at: https://www.gedai.com.br/ghost-writer-uma-breve-analise-a-partir-do-caso-bruna-surfistinha/. (Accessed: 10 august 2023).

Walsh, T. (2022) *Machines behaving badly: The morality of AI.* Chelteham: Flint, versão Kindle.

Zanini, L. E. A. (2015) *Direito de Autor*. São Paulo: Saraiva.