

Sem referências: o ChatGPT sob a perspectiva latouriana do duplo clique

José Victor Rodrigues Catalano¹

Bruno Rossi Lorenzi²

Resumo: A Inteligência Artificial Generativa tem despertado grande interesse na população, e o ChatGPT é um exemplo notável desse fenômeno, devido à sua ampla cobertura midiática. Entretanto, ao analisarmos as principais controvérsias em torno da versão 3.5 do ChatGPT, fica evidente que a ausência de referências confiáveis é um problema alarmante. Em alguns casos, a ferramenta gera informações imprecisas ou falsas, que alguns autores chamam de 'alucinações'. Este artigo foca na problemática das fontes e referências, que o ChatGPT frequentemente não fornece e, em alguns casos, até falsifica. Com a crescente disseminação do uso da ferramenta, a questão da falta de fontes e ou referências pode se tornar ainda mais grave. Os usuários podem considerar o ChatGPT como um oráculo infalível, aumentando sua confiabilidade pelo simples hábito de uso. Nesse sentido, analisamos esse comportamento à luz do conceito de Duplo Clique, presente na obra 'Investigação sobre os modos de existência: Uma antropologia dos modernos', de Bruno Latour. Abordando a ignorância das mediações durante o processo de conversação (saídas) com a IA e seus possíveis efeitos para a sociedade, caso acreditar cegamente em saídas com falta de referências (um duplo clique) se torne um hábito.

Palavras-chaves: ChatGPT, Inteligência Artificial, Bruno Latour, Duplo Clique, Referências.

No references: the ChatGPT from the latourian perspective of the double click

Abstract: *Generative Artificial Intelligence has aroused great interest in the population, and ChatGPT is a notable example of this phenomenon, due to its wide media coverage. However, when analyzing the main controversies surrounding version 3.5 of ChatGPT, it is clear that the lack of reliable references is an alarming problem. In some cases, the tool generates inaccurate or false information, which some authors call 'hallucinations'. This article focuses on the problem of sources and references, which ChatGPT often fails to provide and in some cases even falsifies. As the use of the tool becomes more widespread, the issue of missing sources and or references may become even more serious. Users may consider ChatGPT as an infallible oracle, increasing its reliability by the simple habit of use. In this sense, we analyze this behavior in the light of the concept of Double Click, present in the work 'Investigation into the modes of existence: an anthropology of the modern', by Bruno Latour. Addressing the ignorance of mediations during the process of conversation (outputs) with AI and its possible effects for society, if blindly believing in outputs with lack of references (a double click) becomes a habit.*

Keywords: *ChatGPT, Artificial Intelligence, Bruno Latour, Double Click, References.*

¹ Mestrando no Programa de Pós-Graduação em Ciência, Tecnologia e Sociedade da Universidade Federal de São Carlos (UFSCar). Possui graduação em Administração, pela Universidade Paulista (UNIP). Desenvolve pesquisa na linha das dimensões sociais da ciência e da tecnologia. Orcid: <https://orcid.org/0000-0002-7799-7393> Email: falecomjosevictor@gmail.com

² Pós-doutorando no Programa de Pós-Graduação em Ciência, Tecnologia e Sociedade da Universidade Federal de São Carlos (UFSCar). Possui graduação em Ciências Sociais (2009), mestrado em Ciência, Tecnologia e Sociedade (2012) e doutorado em Ciência Política (2018) pela Universidade Federal de São Carlos (UFSCar). Orcid: <https://orcid.org/0000-0001-8486-0726> Email: brunolorenzi@gmail.com

Sin referencias: el ChatGPT bajo la perspectiva latouriana del doble click

Resumen: *La Inteligencia Artificial Generativa ha despertado un gran interés en la población, y ChatGPT es un ejemplo notable de este fenómeno, debido a su amplia cobertura mediática. Sin embargo, al analizar las principales polémicas que rodean a la versión 3.5 de ChatGPT, queda claro que la falta de referencias fiables es un problema alarmante. En algunos casos, la herramienta genera información inexacta o falsa, que algunos autores denominan "alucinaciones". Este artículo se centra en el problema de las fuentes y referencias, que ChatGPT a menudo no proporciona y, en algunos casos, incluso falsifica. Con la creciente difusión del uso de la herramienta, el problema de la falta de fuentes y/o referencias puede llegar a ser aún más grave. Los usuarios pueden considerar ChatGPT como un oráculo infalible, aumentando su fiabilidad por el simple hábito de uso. En este sentido, analizamos este comportamiento a la luz del concepto de Doble Clic, presente en la obra 'Investigación sobre los modos de existencia: una antropología de lo moderno', de Bruno Latour. Abordando el desconocimiento de las mediaciones durante el proceso de conversación (outputs) con la IA y sus posibles efectos para la sociedad, si creer ciegamente en outputs con falta de referencias (un doble clic) se convierte en un hábito.*

Palabras clave: *ChatGPT, Inteligencia Artificial, Bruno Latour, Doble Clic, Referencias.*

Introdução: abrindo a caixa-preta do ChatGPT

O modelo mais recente da OpenAI, um laboratório de pesquisas em Inteligência Artificial (IA) com sede nos Estados Unidos, está gerando um grande *hype* pela internet. O ChatGPT, é uma IA Generativa que, por meio de um robô conversacional (*chatbot*), emprega técnicas de processamento de linguagem natural e aprendizado profundo. Graças a essas técnicas, ele é capaz de entender e fornecer respostas e soluções a perguntas e solicitações durante uma conversa em linguagem natural. A ferramenta foi disponibilizada em fase de teste em 30 de novembro de 2022, ela é baseada no InstructGPT.3, um modelo da mesma empresa, que tem a capacidade de oferecer explicações naturais para perguntas simples. No entanto, o que o diferencia dos demais modelos é a sua habilidade de interagir de forma contínua com um usuário humano, por meio de feedbacks e entradas. Isso permite que as conversas com o chatbot sejam longas e até mesmo aparentemente intermináveis, até que a interação seja finalizada. (CHATTERJEE; DETHLEFS, 2023).

O debate em torno da IA não é recente, em 1955, durante um seminário de verão na Faculdade de Dartmouth, nos Estados Unidos, John McCarthy criou o termo "Inteligência Artificial". Isso deu início a um movimento de pesquisa e desenvolvimento de sistemas inteligentes que têm como objetivo fazer com que esses sistemas possam pensar e agir como seres humanos. Alan Turing é considerado o pai da Inteligência Artificial (IA) por causa de seu estudo em 1950 sobre a possibilidade de máquinas pensarem. Desde então, a Inteligência Artificial evoluiu como uma área interdisciplinar com o objetivo de criar sistemas que possam executar tarefas complexas que antes só poderiam ser feitas por seres humanos (LUGER, 2004).

Ao longo dos anos de pesquisa, diversos modelos e arquiteturas foram desenvolvidos, porém é o modelo generativo que vem ganhando destaque em publicações jornalísticas, postagens na internet e interações entre usuários em redes sociais. A Inteligência Artificial Generativa é um tipo de rede neural que tem a capacidade de aprender a gerar novos dados realistas, como imagens, sons ou textos, a partir de exemplos fornecidos. Elas funcionam com duas redes neurais: uma rede geradora e uma rede discriminadora, que trabalham juntas para melhorar a capacidade do modelo de gerar exemplos cada vez mais realistas. Essa técnica tem sido aplicada com sucesso em diversas áreas, como a geração de imagens e vídeos, produção de música e criação de texto (COOPER, 2021) (MCKINSEY & COMPANY, 2023). Um exemplo de IA Generativa é o LLM (*Learning-Based Language Model*) ou modelo de linguagem baseado em aprendizado de máquina, é um modelo de IA do campo de processamento de linguagem natural capaz de aprender as relações entre palavras e frases de um idioma a partir de grandes quantidades de dados de treinamento. Com isso, o LLM pode gerar um texto coerente e relevante para responder a perguntas ou executar tarefas específicas.

GPT-3 (*Generative Pre-trained Transformer-3*) é um LLM formado por uma rede neural com cerca de 175 bilhões de parâmetros gerados a partir de mais de 75 terabytes de dados coletados da internet. No entanto, a principal limitação do GPT3 é que o mesmo foi treinado sem supervisão. Por isso, os desenvolvedores do ChatGPT treinaram o GPT3.5, um modelo aprimorado através da adaptação do GPT-3 original em que eles empregaram uma técnica denominada aprendizado por reforço com *feedback* humano visando otimizar o modelo, a fim de garantir que o mesmo não apresente vies e seja capaz de fazer perguntas de acompanhamento quando não estiver confiante em suas respostas. Durante o processo de treinamento, avaliadores humanos forneceram ao modelo uma avaliação positiva (recompensa) quando ele gerava respostas realistas, e uma avaliação negativa (penalidade) quando ele gerava saídas estranhas. Com essa abordagem, o ChatGPT é capaz de combinar as vantagens do GPT anterior, que se baseia no treinamento com grandes quantidades de informações disponíveis na Internet, com a habilidade de manter diálogos que se assemelham ao de um ser humano (BROWN et al, 2020) (CHATTERJEE; DETHLEFS, 2023).

Neste artigo, buscaremos analisar as principais controvérsias em torno da versão 3.5 do ChatGPT, em especial a ausência de referências na maior parte de suas respostas. Em alguns casos, a ferramenta gera informações imprecisas ou até mesmo falsas, que alguns autores chamam de 'alucinações'. Com a crescente disseminação do uso da ferramenta, a questão da falta de fontes e ou referências pode se tornar ainda mais grave. Os usuários podem considerar o ChatGPT como um oráculo infalível, aumentando sua confiabilidade pelo simples hábito de uso. Nesse sentido, analisamos esse comportamento à luz do conceito de Duplo Clique, presente na obra 'Investigação

sobre os modos de existência: Uma antropologia dos modernos', de Bruno Latour. Para este autor, que abordaremos após a exposição das controvérsias relacionadas ao ChatGPT até o momento, o uso de ferramentas que omitem os caminhos e transformações da informação - ou seja, o que a Teoria Ator-Rede (TAR) chama de mediações -, como o ChatGPT nesse caso, pode se tornar uma armadilha (o chamado Duplo Clique) à medida que nos habituamos com as respostas fáceis e imediatas. Ao não trazer as referências e fontes junto às suas respostas, o ChatGPT e outras IAs generativas dificultam o questionamento de suas respostas, mesmo quando erram ou inventam alguma coisa. Sem uma crítica ou regulamentação deste fato, corremos o risco de nos tornarmos em grande medida dependentes de um falso oráculo que, ao contrário das ciências, não nos permite consultar e questionar suas referências. Após essa introdução, abordaremos algumas controvérsias relacionadas ao ChatGPT para em seguida tratar dos conceitos de mediação e de Modos de Existência, em especial o conceito de Duplo Clique e Hábito de Bruno Latour, juntamente com a aplicação destes conceitos à questão da ausência de referências e “alucinações” do ChatGPT. Por fim, trazemos algumas considerações finais sobre os riscos relacionados a esse comportamento em IAs generativas.

ChatGPT entre controvérsias

Em apenas cinco dias após o seu lançamento inaugural, o ChatGPT conseguiu atingir a marca de um milhão de usuários únicos. Enquanto plataformas “gigantes” como a Netflix que precisou de 3,5 anos, o Twitter de 2 anos, o Facebook de 10 meses, o Spotify de 5 meses e o Instagram de 2,5 meses para atingir a marca de um milhão de usuários (KARUNANIDHI, 2023), neste curto período de tempo em relação às demais, tem sido observado um aumento exponencial na adoção do ChatGPT por parte dos usuários. Já que, em apenas dois meses após o seu lançamento, o número de usuários únicos do ChatGPT atingiu cem milhões (HU, 2023). A ferramenta tem se consolidado rapidamente e se tornado objeto de diversas discussões e análises científicas sobre seus múltiplos usos e as controvérsias que o cercam em um curto período de tempo.

O ChatGPT, no âmbito da educação, tem enfrentado restrições consideráveis em várias regiões do mundo. Em cidades como Nova York, Seattle, Los Angeles e Baltimore, nos Estados Unidos, o Departamento de Educação bloqueou o acesso à ferramenta nas redes de internet das escolas públicas, alegando impactos negativos na aprendizagem, devido à preocupação em relação à produção de plágio e desinformação por meio da ferramenta. Adicionalmente, na Austrália, o ChatGPT também foi banido nas redes de internet das escolas públicas em grande parte do país.

Em diversos países, medidas restritivas foram adotadas em relação ao uso do ChatGPT em diferentes níveis da educação. Por exemplo, a Sciences Po, uma renomada universidade francesa, proibiu oficialmente o uso da ferramenta, alegando a falta de transparência em relação às referências. Além disso, a Universidade de Bangalore, na Índia, não só proibiu o uso do ChatGPT, mas também pretende aplicar testes surpresas para detectar estudantes que o utilizam (NOLAN, 2023). Essas medidas refletem a preocupação dos profissionais de diversos níveis no âmbito da educação com o advento da ferramenta na vida acadêmica dos estudantes. Mas, enquanto algumas nações adotaram uma postura cautelosa ou até mesmo belicosa em relação ao uso da ferramenta, Singapura está considerando a possibilidade de integrar o ChatGPT no ensino. O Ministro da Educação fez uma analogia entre o ChatGPT e a introdução da calculadora na educação (outro fato controverso até hoje). Afirmando que a ferramenta pode servir como um auxílio para os estudantes que já possuem conhecimento prévio de conceitos e habilidades de raciocínio (ALVES, 2023). Esse posicionamento destaca a possibilidade de uso do ChatGPT como uma ferramenta educacional, desde que seja utilizado de maneira adequada e ética.

Na escrita acadêmica, a Springer Nature, a maior editora acadêmica do mundo, definiu sua política em relação ao uso de ferramentas de redação de IA em artigos científicos. A empresa afirmou que IAs, como o ChatGPT, não podem ser reconhecidas como autores de artigos publicados em seus periódicos. Contudo, permitirá que cientistas utilizem a IA para auxiliar na escrita ou na geração de ideias para pesquisas, desde que a contribuição da IA seja devidamente mencionada pelos autores (VINCENT, 2023). Por isso, o ChatGPT pode ser útil como um assistente de escrita para ajudar na geração de ideias, sugestões de vocabulário e aprimoramento da gramática em textos científicos. Por exemplo, se um pesquisador estiver tendo dificuldades para encontrar as palavras certas para descrever uma metodologia de pesquisa, ele pode utilizar a ferramenta para sugerir palavras relevantes que possam ser utilizadas.

Como assistente de um advogado, o ChatGPT pode ser utilizado para ajudar na geração de ideias e sugestões de vocabulário para a produção de documentos legais, como petições, contratos e pareceres jurídicos. Além disso, a ferramenta pode ser útil para a elaboração de argumentos em processos judiciais, fornecendo sugestões de palavras-chave e possíveis linhas de raciocínio. Um juiz na Colômbia fez uso da ferramenta para fundamentar sua decisão em um caso específico relacionado à gratuidade do tratamento para uma criança autista. Ele utilizou a mesma para obter informações relevantes para sua decisão, destacando a facilidade que a ferramenta proporcionou para a redação do texto. Em uma entrevista de rádio, o juiz elogiou o ChatGPT, comparando-o a um secretário que o auxiliou no processo de redação do texto (DW, 2023). Um deputado estadual em Santa Catarina também fez uso da ferramenta ao elaborar um projeto de lei que visa promover

a transparência do estoque de medicamentos na rede de saúde pública estadual (FREITAS, 2023). Especialistas jurídicos afirmam que o ChatGPT pode ser uma ferramenta útil para auxiliar na produção jurídica. No entanto, é importante destacar que a ferramenta não deve ser vista como uma solução definitiva para a produção de textos jurídicos. Pois, cabe ao advogado adequar o texto gerado pelo ChatGPT ao contexto específico do caso em questão, levando em consideração a legislação aplicável e as necessidades do cliente. Além disso, é essencial que o advogado realize uma revisão crítica do texto gerado, a fim de garantir que as informações apresentadas estejam corretas e coerentes com a argumentação proposta.

Portanto, o uso do ChatGPT como facilitador na produção de documentos jurídicos deve ser feito com cautela e em conjunto com o conhecimento e habilidades do advogado. A ferramenta pode ser uma aliada na otimização do tempo e dos recursos utilizados no processo de produção, mas não deve substituir a expertise e a análise crítica do profissional de direito (MIGALHAS, 2023a). Esta que foi “aprovada” por meio de estudos científicos realizados em exames de qualificação em áreas como medicina, advocacia e MBA na renomada Universidade de Wharton (SILVA, 2023). No Brasil, ela também obteve aprovação na primeira fase do exame da Ordem dos Advogados do Brasil (MIGALHAS, 2023b).

No que tange o criativo, estima-se que a ferramenta já criou ao menos 200 livros que estão sendo vendidos na Amazon (SÉRVIO, 2023). Um escritor em especial, utilizou o ChatGPT e outra ferramenta denominada Midjourney (uma IA generativa de imagens) para publicar um livro infantil em menos de 72 horas (FERREIRA, 2023). Os desenvolvedores de jogos também estão utilizando o ChatGPT, para auxiliar os seus processos criativos e otimizar a produção de títulos, principalmente na área de narrativas. Ao utilizar o ChatGPT para gerar ideias e sugestões de enredos, os desenvolvedores economizam tempo e recursos durante o processo de desenvolvimento do jogo (SYOZI, 2023a).

Na seara da informática em geral, a ferramenta foi aprovada em uma entrevista de emprego na área de codificação no Google (VENTURA, 2023). O CEO da Nvidia, empresa multinacional de tecnologia de processadores gráficos, acredita na "democratização" da programação por meio da ferramenta, pois a mesma permite receber perguntas em linguagem humana e traduzir para linguagens de programação (SYOZI, 2023b). Isso, trouxe preocupação para os programadores, na qual Welsh (2023) em um artigo denominado "The End of Programming" (O Fim da Programação), acredita que a programação, principalmente a "escrita de programas" se tornará obsoleta, já que assistentes como o CoPilot (uma IA generativa da plataforma GitHub) e até mesmo o GPT3 da própria OpenIA possuem potencial para substituir os programadores.

Desde o seu lançamento há três meses, o ChatGPT tem sido objeto de discussão em diversos setores, abrangendo áreas antagônicas, como educação, criatividade, direito e economia. Sua crescente influência chamou a atenção da mídia, tendo sido destaque na capa da revista Time em fevereiro de 2023 (CHOW; PERRIGO, 2023). Será que o ChatGPT será considerado a personalidade do ano de 2023?

Em quesito de personalidade, um fenômeno que também está gerando controvérsias está relacionado à “alucinações”. Uma alucinação em IA ocorre quando a inteligência artificial fornece uma resposta confiável, mas que não pode ser explicada pelos dados de treinamento, levando-a a fazer afirmações independentes. De acordo com Ji et al. (2022), o conteúdo gerado pelas alucinações é geralmente sem sentido ou incorreto e pode ser categorizado em alucinações intrínsecas e extrínsecas. Nas alucinações intrínsecas, a saída (resposta) gerada que entra em contradição com o conteúdo original. Como um exemplo citado por Ji, et al (2022, p.4) ao perguntar sobre a aprovação da vacina contra o Ebola: "O resumo gerado “A primeira vacina contra o Ebola foi aprovada em 2021” contradiz o conteúdo de origem “A primeira vacina para o Ebola foi aprovada pelo FDA em 2019. ”. Já as alucinações extrínsecas referem-se à geração de saída que não pode ser verificada a partir do conteúdo da fonte original. Pois a saída gerada a partir da fonte não possui evidências que comprovem sua veracidade, e tampouco podemos afirmar que esteja incorreta. É importante notar que a alucinação extrínseca pode não ser necessariamente incorreta, já que pode incluir informações externas que sejam factualmente corretas. No entanto, a maioria dos estudos segundo Ji, et al (2022, p.4) aborda a alucinação extrínseca com cautela devido à sua natureza não verificável, o que aumenta o risco em termos de segurança factual (JI et al., 2023).

Como exemplo, no estudo conduzido por Cortiz (2023), em que o autor solicitou ao ChatGPT uma lista de livros de sua orientadora de doutorado, a semioticista Lúcia Santaella. Em vez de receber uma lista precisa e confiável, o autor recebeu uma série de títulos gerados pela inteligência artificial com base em palavras relacionadas aos temas de pesquisa da professora. Hillier (2023) em sua pesquisa, revelou que 5 das 6 referências geradas pelo ChatGPT eram falsas, explicando que as referências falsas são criadas a partir de elementos que parecem plausíveis pelo modelo estatístico do GPT3, que resulta na perda de informações relevantes e faz com que o modelo crie combinações plausíveis. Nomes reais de autores, periódicos e títulos podem ser usados, mas geralmente são combinados de maneira incorreta ou falsa. Além disso, os intervalos de datas e números de página também são frequentemente imprecisos (HILLIER, 2023).

Pesquisadores das universidades de Northwestern e de Chicago conduziram um estudo em que orientaram o ChatGPT a criar resumos fictícios de artigos publicados em revistas de saúde. Esses resumos foram submetidos a um estudo cego, no qual revisores humanos avaliaram

documentos falsos e reais e foram solicitados a distinguir entre eles. Os resultados mostraram que os revisores conseguiram identificar corretamente 68% dos resumos gerados pelo ChatGPT, mas confundiram 14% dos resumos originais como sendo gerados pela IA. Os revisores notaram que os resumos gerados pela IA eram mais genéricos e tinham uma sensação estereotipada à escrita, apesar de que alguns revisores acharem uma tarefa surpreendentemente difícil em determinar quais dos resumos eram reais e quais eram gerados pelo ChatGPT (GAO et al, 2022).

ChatGPT: uma nova forma de Duplo Clique?

Bruno Latour (1947-2022) foi um importante filósofo, antropólogo e sociólogo do período recente. Debruçou-se sobre vários temas ao longo de sua carreira, especialmente ciência e tecnologia, meio ambiente, democracia e epistemologia. O autor é considerado um dos maiores expoentes da sociologia pragmática francesa e um dos fundadores da Teoria Ator-Rede. Conhecido por obras famosas, que destoam do pensamento sociológico e epistemológico convencional, como *Jamais Fomos Modernos* (1994) e mais recentemente *Investigação Sobre os Modos de Existência* (2019), o autor enfatiza noções e princípios que se inspiram em muitos ramos da teoria do conhecimento, desde a sociologia tardiana, passando pela filosofia do processo de Alfred Whitehead e autores mais contemporâneos como Michel Serres, Gilles Deleuze e Félix Guattari, entre muitos outros (HARMAN, 2009). Em seus estudos de caso, sempre muito próximos ao campo e ao conteúdo científico, tentou abrir a caixa-preta da ciência e da tecnologia, buscando demonstrar como era complexo e heterogêneo esse “interior”, muito mais conectado com a parte de fora do que imaginávamos, além de sempre apontar a importância dos atores não humanos, muitas vezes relegados nas análises sociológicas mais tradicionais.

Em sua última obra de caráter mais teórico, “Investigação Sobre os Modos de Existência: uma antropologia dos Modernos” (*IME*), Latour (2019) faz um panorama do alcance e das limitações da TAR, colocando-a em paralelo a outras formas de análise e de atuação (ou passar a existir/desaparecer), que ele vai chamar de Modos de Existência.

Apesar de não refutar a TAR, Latour a posiciona, agora, como um dos seus quinze modos de existência, aproximando-a ao que se entende por epistemologia e, talvez mais ainda, por metodologia, mas que na linguagem latouriana aparece como “metalinguagem da investigação”. Já no primeiro capítulo de *IME*, Latour apresenta esta metalinguagem da investigação, que antes chamava de TAR, como *Modo Rede*, com grafia [RES] (entre colchetes, mesmo) para designá-lo como parte integrante do vocabulário técnico adotado.

Ao caracterizar o modo [RES], Latour (2019, p. 39–46) explica que se trata da arte de explorar entidades necessárias à existência de outras entidades, já que redes são coletivos de relações, composições de elementos heterogêneos articulados de forma a produzir um sistema relacional dinâmico e instável, mas que, devido às ações de manutenção, aparenta ser estável. Por tratar todo coletivo como associação de heterogêneos, [RES] seria uma espécie de método de investigação similar ao de um antropólogo, isto é, um método interessado em entender uma cultura (próximo à etnometodologia), que adota uma postura conhecida como um ceticismo metódico (que propõe uma ida a campo a partir de um não saber) e de base empírica (já que propõe um descrever dos elementos que agem no coletivo, a partir daquilo que a experiência apresenta como fenômeno), e que sugere uma espécie de achatamento do social em questão, destacando, a cada etapa de descrição, seu arranjo composicional. No entanto, partir da rede não significa empreender uma busca totalmente errática e desinteressada. O que se pretende é ressaltar os elementos que agem em um coletivo, visando captar a construção de sua episteme.

Segundo o autor de *IME*, é preciso perceber como diferentes modos de existência coexistem e se cruzam se quisermos entender como produzimos continuidades (ou seja, sentidos, sejam eles na forma de conhecimento, tecnologia, leis, etc.), a partir das discontinuidades que permeiam o mundo. Em suas palavras:

[...] aprendemos nos capítulos anteriores que nossa investigação tinha a ver com a identificação de um tipo de TRAJETÓRIAS cuja continuidade aparente era de fato obtida por uma maneira peculiar de saltar sobre as discontinuidades, que são diferentes a cada ocasião (LATOURE, 2019, p. 69-70).

De acordo com o autor, cada modo de existência possui a sua chave de leitura, ou Preposição [PRE], assim como suas condições particulares de felicidade e infelicidade (que podemos interpretar como suas condições de sucesso). Latour (2019) aponta quinze modos distintos de existência, sem a intenção de esgotar as possibilidades. Neste artigo não iremos explorar todos esses modos, mas é importante citar o modo de existência, por excelência, do conhecimento científico.

Segundo Latour (2019), modo Referência [REF] é a forma pela qual nós encontramos de conectar coisas diferentes por meio de suas semelhanças, categorizadas a partir do nosso entendimento do mundo. Esse modo encontra sua versão mais desenvolvida no que chamamos de ciências (naturais, sociais ou do direito), já que essas atividades seriam impraticáveis, ou não teriam nenhuma legitimidade, não fossem as longas cadeias de referência que criam, possibilitando o resgate da informação, a revisão por pares e o encadeamento lógico de modelos, conceitos e

teorias científicas. Mas ao contrário do direito, que sobrevive perfeitamente encadeando apenas referências legais e morais, a ciência precisa das entidades físicas da Reprodução (antes chamados de actantes não humanos) para construir seus experimentos, produtos e artefatos. É, portanto, a partir do cruzamento de [REP] e [REF] que é possível, por exemplo, desenvolver um novo material para a indústria, criar um fármaco, manipular geneticamente um organismo, etc. Nas palavras do autor:

Vamos, portanto, denominar como [REP] a REPRODUÇÃO (insistindo, é claro, no prefixo "re", de reprodução), o modo de existência pelo qual uma entidade qualquer atravessa o hiato de sua repetição, definindo assim, de etapa em etapa, uma trajetória particular, em que o conjunto obedece a condições de felicidade particularmente exigentes: ser ou não ser mais! Sem surpresas, usaremos a notação [REF] (de REFERÊNCIA) para designar o estabelecimento das cadeias definidas pelo hiato entre as duas formas de natureza diferente e cuja condição de felicidade consiste na descoberta de uma constante que se mantém através desses abismos sucessivos, desenhando outra forma de trajetória que torna acessíveis os distantes, cobrindo o trajeto com o movimento em duplo sentido dos móveis imutáveis (LATOURET, 2019, p. 85).

Para o autor, é perfeitamente possível rastrear o cruzamento entre as cadeias de referência das ciências [REF] e os actantes não humanos a elas associados [REP] a partir da TAR - ou, para utilizar os termos da obra, o modo [RES] -, já que essas associações se dão na forma de redes heterogêneas, cujas atuações e resistências deixam rastros visíveis ao analista na forma de inscrições (como os artigos, patentes e discussões acadêmicas), as vezes ainda mais notáveis na forma de controvérsias, o que geralmente abre uma oportunidade para esmiuçarmos até seus mais ínfimos e ocultos detalhes.

No entanto, isso só nos é possível porque o modo [REF] possui como fundamento (ou o que o autor vai chamar de “condições de felicidades”) o apontamento rigoroso da fonte ou origem das informações. Nenhuma afirmação, portanto, é feita sem a devida referência ou apresentação dos dados em que foi baseada. Isso é justamente o que distingue o conhecimento científico das simples opiniões ou crenças. Ao trazer em si mesmo a cadeia de referências que deram origem às conclusões, um artigo científico ou obra de divulgação científica, assim como uma aula devidamente preparada, pavimenta o percurso dessas inscrições, trazendo legitimidade e solidez ao argumento.

É importante abordarmos brevemente este que é, provavelmente, o conceito central da Teoria Ator-Rede e na epistemologia latouriana como um todo, o conceito de translação, também chamado de mediação. É preciso ressaltar aqui, para evitar confusões, que o autor usa estes termos de forma completamente intercambiável, dando preferência a um ou outro de acordo com a obra ou situação. Entretanto, se nos aprofundarmos na leitura do conjunto da obra, perceberemos que

estes termos têm significados muito próximos e complementares, especialmente a partir do uso que o autor (e demais autores da TAR, como Michel Callon e John Law) faz deles. Isso significa que, de modo geral, as ações, ou melhor dizendo, as atuações de qualquer actante (humano ou não humano), no entendimento da TAR, são sempre ao mesmo tempo: *traduções*, já que implicam modificações e interpretações da realidade; *translações*, posto que envolvem deslocamentos de outros actantes em torno de algum ponto ou objetivo e; *mediações*, já que sempre é necessário utilizar ou se criar novas conexões – ou *associações* – entre actantes para se realizar uma ação. Nas palavras do autor: “empreguei translação para indicar deslocamento, tendência, invenção, mediação, criação de um vínculo que não existia e que, até certo ponto, modifica os dois originais” (Latour 2001, 206), ou ainda, “em lugar de uma rígida oposição entre contexto e conteúdo, as cadeias de translação referem-se ao trabalho graças ao qual os atores modificam, deslocam e transladam seus vários e contraditórios interesses” (Latour 2001, p. 356).

Ao focar seu método de análise no mapeamento das translações-mediações, a TAR buscou identificar - a partir do relato dos próprios actantes ou através das controvérsias que desnudam um fato ou situação ainda não estabilizada - o caminho ou trajetória que os atores-redes trilharam para realizar seus objetivos, ou seja, as alianças com atores humanos ou não humanos que precisam fazer para tanto, as estratégias que utilizam para seduzir esses atores a participarem dessa rede, as modificações que esses atores-mediadores sofrem e produzem nesse processo, as diferentes interpretações da realidade que atores diferentes da rede ou concorrentes trazem sobre um assunto ou fato, etc. Dessa forma, segundo Latour (2012), é possível investigar o que realmente ocorre na produção científica, tecnológica ou mesmo política sem apoiar em apriorismos sociológicos (como estrutura, luta de classes, fatos sociais, tec.) que atuam como uma dimensão extra da realidade, explicando o que na verdade deveria ser explicado. São justamente esses movimentos que implicam necessariamente em transformações (traduções) que o autor buscou, ao longo de sua obra, explicitar e explicar, num esforço de ser acuradamente descritivo e imparcial, ou ao menos sem a “arrogância sociológica” de, supostamente, saber mais sobre um assunto do que os próprios informantes, ainda mais quando os assuntos preferidos ao longo de sua obra foram as ciências e a tecnologia.

A partir desse conceito, que na TAR atua como um princípio, torna-se possível para o analista produzir análises detalhadas e acuradas sobre praticamente qualquer assunto, já que podemos ser leigos sobre o assunto que estudamos e ainda assim produzir um panorama ou mapa da rede rico o suficiente para contribuir com o seu entendimento, desde que tomemos os cuidados ressaltados pelo autor (Latour, 2012), como não realizar “saltos” ou pegar “atalhos” em conceitos generalizantes. Entretanto, fica muito difícil rastrear os atores envolvidos em alguns fenômenos

ou acontecimento, suas estratégias, seus movimentos e suas transformações, quando essa trajetória é propositalmente ocultada.

Latour (2019) alerta sobre os perigos de apagarmos a trajetória das inscrições e dos seres que fazem parte dessa rede (suas mediações) e pularmos direto para a resposta ou resultado, muitas vezes de forma parcial, distorcida ou enganosa, ao buscar conectar duas coisas sem citar ou descrever as traduções, ou seja, as transformações que essas entidades sofreram ao longo do caminho até se estabilizarem. É o que ele vai chamar de perigo do Duplo Clique, em alusão à facilidade que os dispositivos conectados à internet nos trouxeram nas últimas décadas quanto ao acesso à informação.

Latour caracteriza o modo Duplo Clique como uma espécie de gênio maligno que, ao pretender falar reto, ir direto ao assunto, sem fazer desvios ou enrolações, conecta duas ou mais coisas de forma muito direta e dessa forma de difícil contestação. O objetivo aqui é justamente apagar as mediações, fazer saltos sem pagar o preço das traduções, ou seja, sem considerar as transformações de várias naturezas que os seres sofrem ao longo do processo, ocultando os desvios e a trajetória de cada modo específico, e assim, essencializando ou naturalizando as coisas (Latour, 2019, p. 160). As chamadas *fake news* utilizariam esse modo por excelência, já que necessitam apagar todos os fundamentos e caminhos tortuosos de um fato (verdadeiro ou não) para tornar o debate impossível. Ao mesmo tempo, elas aparentam serem mais verídicas, justamente porque as pessoas não estão acostumadas com os caminhos tortuosos e controverso da ciência e da política, que envolve dúvidas, debates, incertezas, diferentes visões e opiniões sobre um mesmo assunto - por mais que seja científico.

Ainda que seja mais fácil obter informações por meio de nossos computadores e smartphones, com um simples clique ou comando de voz, a maior parte das vezes essa informação nos é trazida sem citar suas fontes ou trajetória, num caminho oposto ao das ciências. Consequentemente, nem sempre essas informações precisam ser verdadeiras ou trazerem consigo seus fundamentos. Para o autor, essa é a principal ameaça que os analistas, assim como os diplomatas da política, da ciência e dos movimentos sociais, devem enfrentar:

Em alusão ao mouse do computador, vamos chamar esse diabo de Duplo Clique (e representá-lo pela notação [DC]). A partir de uma experiência perfeitamente exata - a referência permite o acesso - esse Gênio Maligno sussurrará em nosso ouvido que seria muito melhor termos acesso gratuito, indiscutível e imediato à informação pura e sem transformação. Agora, se, infelizmente, esse ideal de gratuidade total servisse de padrão para julgar o verdadeiro e o falso, então tudo se tornaria mentiroso, inclusive a ciência (Latour, 2019, p.136-137).

Até o momento em que escrevemos esse artigo, o ChatGPT não trazia junto às suas respostas basicamente nenhuma fonte ou origem das informações, ou pior, às vezes as inventa ou alucina!, funcionando basicamente como um oráculo para quem o utiliza. Ao conectar a pergunta do usuário à resposta de forma direta, sem apresentar o caminho da informação e suas fontes (ou seja, as mediações ou referências do modo [REF], tão caras às ciências) o ChatGPT (e outras IA generativas semelhantes) pode ser caracterizado, na teoria latouriana, como uma forma de Duplo Clique. A facilidade e rapidez com que uma IA generativa desse tipo nos traz uma resposta, sobre praticamente qualquer assunto, nos acende muitos alertas, que vão do uso indevido de direitos autorais e não citação de fontes até a possibilidade de respostas erradas de difícil identificação, já que o caminho até o resultado não nos é apresentado.

Podemos dizer que o ChatGPT funciona de forma completamente inversa ao modo [RES] (ou Modo Rede), ocultando do usuário a trajetórias das mediações e, por consequência, todo o debate e disputas que envolvem qualquer posição científica ou política. É necessário, portanto, que sejamos cautelosos com as IAs generativas, posto que elas são desenvolvidas e treinadas a partir uma base imensa de dados, muito além das capacidades humanas, mas nos trazem resultados que podem conter violações ou imprecisões de difícil identificação por pessoas não treinadas e mesmo especialistas no assunto.

Segundo Latour (2019, p. 357-358) o Hábito (ou [HAB]) é um modo de existência cuja a função é justamente omitir, ocultar as demais preposições que dão o sentido de uma trajetória e atuam como sinalizadores dos modos de existência. Por exemplo, quando assistimos à um filme e nos esquecemos que é uma obra de ficção, a ponto de nos emocionarmos [HAB x FIC], quando um electricista atua com tanta confiança que não precisa nem desligar a chave geral [HAB x TEC], ou mesmo quando nem nos preocupamos em conferir todas as referências citadas em um artigo, tamanha nossa confiança naquele cientista ou autor [HAB x REF]. O hábito nos ajuda em basicamente todas as tarefas do dia, permitindo a prática sem despender muita energia pensando naquilo. Porém, tem como efeito colateral tornar muito dispendioso e desafiador a dúvida e o questionamento de coisas que estão tão bem estabelecidas na nossa vida.

Segundo Latour, o perigo se dá justamente quando o Hábito cruza com o Duplo Clique ([DC x HAB]). Ao nos habituarmos com as facilidades que a internet ou agora as IAs generativas nos proporcionam, corremos o risco de esquecer que na realidade muitas vezes são apenas atalhos, e atalhos nem sempre nos levam ao destino esperado. “Uma vez inventado o padrão de uma palavra reta, todos os outros passam, em comparação, a falar atravessado; eles se tornam pífidos, mentirosos, manipuladores” (Latour, 2019, p.181). Vimos um fenômeno semelhante com as redes sociais, já que muitas pessoas, na maioria das vezes fechadas em suas “bolhas”, dificilmente

duvidam de algo dito por pessoas próximas ou de sua confiança, dispensando qualquer fundamentação sobre algum pretensão “fato” ou posicionamento. O mesmo pode acontecer com as IA generativas como o ChatGPT. À medida que essa ferramenta é aperfeiçoada e nos habituamos a utilizá-la para praticamente tudo, ficará difícil identificar erros ou mesmo questioná-los quando encontrados, já que as pessoas podem se acostumar com a facilidade e a credibilidade “infalível” da rede neural. Um caminho sem saltos, sem transformações, direto ao ponto, torna-se sedutor. E o que fazer se esse sistema, por ventura, tornar-se impreciso, parcial ou mesmo manipulativo - de acordo com o interesse de seus mediadores (financiadores, acionistas, desenvolvedores, diretores, etc)? Como o autor alerta:

Porque se a crise surge (...) então estamos realmente perdidos, incapazes de reparar, de retomar, de reencontrar as ramificações perdidas. O que era apenas um legítimo e ligeiro velamento, uma necessária omissão, foi transformado em esquecimento. Não há mais "reinicialização manual". Sem retomada, é a catástrofe, no avião só há um piloto automático. Conseguiremos salvar o Duplo Clique de si mesmo? Torná-lo sensível à quantidade vertiginosa de mediações para que um clique de mouse produza um efeito qualquer? Podemos reconciliá-lo com a sua verdadeira etologia, aquela dos melhores de linhas de código que foi preciso escrever com grandes despesas para que um duplo clique conseguisse efetivamente produzir um efeito? (Latour, 2019, p. 373-374).

E aqui chegamos ao ponto crítico de nossos apontamentos. O ChatGPT vem ocupando cada vez mais espaços e funções no nosso dia-a-dia, e a tendência é que isso cresça cada vez mais rápido. Montagem de processos jurídicos e administrativos, redação de contos e notícias, organização de eventos e processos, programação de softwares, desenhos de engenharia, etc. O ChatGPT demonstrou, até agora, facilidade para realizar todas essas atividades, entre outras, que careciam de ampla experiência e especialização (e com isso custo), e agora se tornam acessíveis e gratuitas. O ChatGPT não apenas simula uma conversa (apesar de esta ser sua interface por excelência, pensado para ter essa faceta desde o início), ele e outras IAs em desenvolvimento se tornaram cada vez mais ferramentas para a execução das mais diferentes e complexas tarefas. Não deve demorar muito para essas ferramentas serem utilizadas também na redação de artigos científicos e como fontes de pesquisa. É por isso que as referências, tão caras às ciências e à TAR, são de suma importância. A ausência de referências não pode se transformar em um hábito, habilitando o gênio maligno do Duplo Clique a atuar em espaços onde a rastreabilidade da informação é fundamental. É, portanto, fundamental discutirmos esse fenômeno sob esse aspecto e talvez propor uma regulamentação antes que seja tarde demais.

Considerações Finais

A Inteligência Artificial (IA) tem sido objeto de pesquisa desde a década de 1950, porém, nos últimos anos, avanços significativos têm sido alcançados em áreas como processamento de linguagem natural e aprendizado profundo, o que tem impulsionado o desenvolvimento de sistemas autônomos capazes de desempenhar tarefas complexas e interagir com seres humanos de forma cada vez mais natural. O ChatGPT é um exemplo notável de progresso na área da IA em direção à meta central de desenvolver sistemas autônomos altamente avançados. A capacidade de manter diálogos em linguagem natural com seres humanos é uma das principais características do ChatGPT, o que reflete a evolução contínua e a sofisticação crescente da tecnologia. Os avanços na IA estão abrindo caminho para novas possibilidades em diversas áreas que antes eram inimagináveis. Isso é resultado de anos de pesquisa e desenvolvimento contínuo, que levaram à criação de sistemas cada vez mais sofisticados e capazes de desempenhar tarefas complexas. Como resultado, a IA está mudando rapidamente a forma como as pessoas interagem com a tecnologia e entre si, com implicações significativas para o futuro da sociedade.

A concepção de que o ChatGPT está transformando as relações entre seres humanos e não humanos é de grande importância, visto que uma mudança significativa está ocorrendo em nossa forma de interação com a tecnologia. Estamos presenciando uma transformação fundamental na maneira como realizamos tarefas que antes eram exclusivamente executadas por humanos, o que nos leva a repensar nossas relações com as máquinas. Com esta ferramenta, temos a capacidade de interagir com uma IA de maneira mais natural e humana do que nunca antes, permitindo que a tecnologia se torne ainda mais integrada em nossas vidas. Isso apresenta o potencial de "reagregar o social", a ponto de a tecnologia mediar relações sociais.

Com sua rápida adoção pelos usuários, é evidente que há um interesse crescente em tecnologias que permitam interações mais naturais e fluidas com a IA. No entanto, é importante reconhecer que o desenvolvimento de modelos de IA não é isento de desafios e limitações. Um dos principais desafios que a IA enfrenta é a questão da ética e do viés nos modelos. Os algoritmos de aprendizado de máquina são treinados com base em conjuntos de dados, e se esses dados forem tendenciosos, o modelo resultante pode perpetuar esses preconceitos. Isso pode ter implicações negativas significativas em diversas áreas, desde a discriminação em sistemas de recrutamento até o viés racial em sistemas de justiça criminal. Outro desafio a ser enfrentado é a questão da privacidade dos dados. À medida que a Inteligência Artificial se torna cada vez mais integrada em nossas vidas cotidianas, é fundamental garantir que os dados pessoais sejam protegidos e que as

informações coletadas não sejam mal utilizadas ou compartilhadas sem o consentimento adequado. Esse aspecto é particularmente crítico em áreas como a saúde, onde a privacidade das informações médicas dos pacientes deve ser preservada.

Em relação ao ChatGPT, um dos principais problemas reside na ausência de referências confiáveis, o que pode resultar em informações imprecisas ou mesmo falsas. Nesse quesito, este artigo buscou dar enfoque ao problema das fontes e referências, que o ChatGPT insiste em não informar ou mesmo falsificar (alucinar). Ao trazer resultados de forma tão rápida e completa, por mais complexo que seja a pergunta ou pedido do usuário, a IA generativa acaba por realmente parecer “mágica” aos olhos da maior parte das pessoas que a utiliza. Dado o crescimento vertiginoso do uso dessa tecnologia que pudemos perceber logo nos primeiros meses de sua estréia, quando este artigo foi escrito, é fácil pressupor que muito em breve o ChatGPT e IAs generativas similares estarão sendo utilizados de forma ainda mais ampla e massiva para os mais diversos fins. É muito provável que o uso de IAs generativas para fins técnicos, acadêmicos e científicos torne-se um hábito em pouco tempo, assim como aconteceu com a internet e as redes sociais.

Porém, na medida em que seu uso se transforme em um hábito por mais e mais pessoas, o problema da ausência de fontes e referências, que possam ser rastreadas e conferidas para checagem dos fatos, gera um problema que tende a se agravar. O risco é justamente que a maior parte dos usuários considerem o ChatGPT uma forma de “oráculo”, que tudo sabe e responde com perfeição. É a armadilha do Duplo Clique que Latour nos alertou. Ao ocultar as fontes e referências, a IA generativa esconde todos os caminhos que a levou até a resposta, tornando quase impossível a checagem dos fatos. Esse cenário tende a se tornar ainda mais preocupante na medida em que a confiança das pessoas no sistema crescer por força do hábito [HAB x DC]. Teremos coragem de questionar a resposta dada pelo robô, de forma tão rápida e eficiente, quando um erro ou imprecisão for identificado? E quem acreditará nos questionadores? Podemos confiar cegamente em suas respostas? Ou seria melhor regulamentar esse tipo de IA generativa para trazer suas fontes junto às respostas enquanto ainda é tempo?

Até o momento, o processo de aprendizado do ChatGPT está sendo supervisionado, de uma forma surpreendentemente responsável. Mas não temos como ter certeza de que esse sistema continuará sendo supervisionado ou o grau de comprometimento ético de seus mediadores humanos. Outras formas de IA generativas, tão ou mais eficazes que o ChatGPT, podem vir a tomar o seu lugar no futuro. Se estas mesmas plataformas não forem tão responsáveis, suas respostas rápidas sobre quase qualquer coisa trazem junto consigo um risco ainda maior se o hábito de as utilizar sem questionamentos já tiver sido pavimentado por outra IA anterior mais confiável.

Portanto, o risco não se encerra no ChatGPT e sua atual ausência de fontes e falsificação de referências. O risco maior é justamente nos habituarmos com o Duplo Clique proporcionado por essas tecnologias e nos esquecermos completamente que elas foram criadas e moldadas por sujeitos com interesses econômicos, políticos e pessoais. Esse fator pode ser especialmente preocupante em áreas como a saúde ou a política, nas quais informações incorretas podem gerar consequências graves. Em um cenário bastante pessimista, as IAs generativas que omitem as fontes e referências - cujos parâmetros são estabelecidos por atores demasiadamente humanos - podem nos levar a crer no que bem quiserem ou lhes for mais conveniente, num processo semelhante ao que vimos acontecer com o fenômeno das *fake news*, divulgadas em massa nas redes sociais, inclusive por figuras bastante populares e até presidentes, e que ainda estamos lutando para regulamentar.

Assim, é necessário que sejam estabelecidos padrões mais rigorosos para a inclusão de referências e fontes confiáveis no ChatGPT, a fim de evitar a disseminação de informações imprecisas ou falsas. Além disso, é preciso que os usuários desenvolvam uma postura mais crítica em relação às informações obtidas por meio da IA generativa, reconhecendo que essas informações são produzidas por sistemas complexos que devem ser avaliados cuidadosamente antes de serem aceitos como verdadeiros. Somente assim poderemos garantir que a IA generativa possa ser utilizada de forma responsável e eficaz, sem comprometer a integridade da informação e a confiança do público nos sistemas de inteligência artificial.

Pois, é essencial que o desenvolvimento das tecnologias de IA seja acompanhado de uma reflexão crítica sobre seus impactos e implicações em nossa sociedade. As tecnologias devem ser socialmente apropriadas de modo a oferecer mudanças que causem um impacto positivo em escala para a sociedade, capazes de reduzir as desigualdades e minimizar nossos erros, com o objetivo de melhorar nossa qualidade de vida. Isso requer um diálogo aberto e colaborativo entre desenvolvedores, pesquisadores, legisladores e o público em geral. Devemos considerar cuidadosamente as implicações de longo prazo da IA em relação à privacidade dos dados, ética e justiça, para que possamos garantir que as tecnologias de IA sejam desenvolvidas e usadas de maneira responsável e benéfica para a sociedade como um todo.

Referências

ALVES, Paula. ChatGPT pode ser usado como “calculadora” na educação, mas com alguns cuidados. **Tecnoblog**, 22 fev. 2023. Disponível em: <https://tecnoblog.net/noticias/2023/02/22/chatgpt-pode-ser-usado-como-calculadora-na-educacao-mas-com-alguns-cuidados/amp/>. Acesso em: 5 mar. 2023.

BROWN, B. MANN, N. RYDER, M. SUBBIAH, J. D. KAPLAN, et al. Language models are few-shot learners. In **Advances in Neural Information Processing Systems**, volume 33, pages 1877–1901, 2020.

CHATTERJEE, J; DETHLEFS, N. This new conversational AI model can be your friend, philosopher, and guide... and even your worst enemy. **Patterns**, New York, N.Y., v. 4, n. 100676, ed. 1, 13 jan. 2023. Disponível em: [https://www.cell.com/patterns/fulltext/S2666-3899\(22\)00323-3](https://www.cell.com/patterns/fulltext/S2666-3899(22)00323-3). Acesso em: 14 fev. 2023.

CHOW, ANDREW R.; PERRIGO, BILLY. The AI Arms Race Is Changing Everything. **TIME**, 16 fev. 2023. Disponível em: <https://time.com/6255952/ai-impact-chatgpt-microsoft-google/>. Acesso em: 7 mar. 2023.

COOPER, K. OpenAI GPT-3: Everything You Need to Know. In: **Springboard Blog**, 1 nov. 2021. Disponível em: <https://www.springboard.com/blog/data-science/machine-learning-gpt-3-open-ai/>. Acesso em: 24 fev. 2023.

CORTIZ, Diogo. Entendendo as Alucinações do ChatGPT. **Diogo Cortiz**, 15 fev. 2023. Disponível em: <https://diogocortiz.com.br/entendendo-as-alucinacoes-do-chatgpt/>. Acesso em: 27 mar. 2023.

DW. Juiz usa ChatGPT para proferir decisão em corte na Colômbia. **DW**, 3 fev. 2023. Disponível em: <https://www.dw.com/pt-br/juiz-usa-chatgpt-para-proferir-decis%C3%A3o-em-julgamento-na-col%C3%B4mbia/a-64602023>. Acesso em: 7 mar. 2023.

FERREIRA, Octávio. ChatGPT: Homem utiliza IA para escrever e publicar livro em menos de 72h; entenda: Até onde irão as IAs?. **IGN**, 17 jan. 2023. Disponível em: <https://br.ign.com/tech-3/105618/news/chatgpt-homem-utiliza-ia-para-escrever-e-publicar-livro-em-menos-de-72h-entenda>. Acesso em: 7 mar. 2023.

FREITAS, Felipe. Deputado estadual de Santa Catarina usa ChatGPT para criar projeto de lei.

Tecnoblog, 2 mar. 2023. Disponível em: <https://tecnoblog.net/noticias/2023/03/02/deputado-estadual-de-santa-catarina-usa-chatgpt-para-criar-projeto-de-lei/amp/>. Acesso em: 7 mar. 2023.

GAO, Catherine A.; HOWARD, Frederick M.; MARKOV, Nikolay S.; YER, Emma C.; RAMESH, Siddhi; LUO, Yuan; PEARSON, Alexander T. Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers. **BioRxiv**, 27 dez. 2022. Disponível em: <https://www.biorxiv.org/content/10.1101/2022.12.23.521610v1>. Acesso em: 18 abr. 2023.

HARMAN, Graham. 2009. **Prince of networks**: Bruno Latour and metaphysics. Melbourne: Re.press. Disponível em: http://www.re-press.org/books/OA_Version_780980544060_Prince_of_Networks.pdf> acesso em: 14/02/2022.

HILLIER, Mathew. Why does ChatGPT generate fake references. **TECHE**. Disponível em: <https://teche.mq.edu.au/2023/02/why-does-chatgpt-generate-fake-references>. Acesso em: 09/03/2023.

HU, Krystal. ChatGPT sets record for fastest-growing user base - analyst note. **Reuters**, 2 fev. 2023. Technology,. Disponível em: <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>. Acesso em: 5 mar. 2023.

JI, Ziwei; LEE, Nayeon; FRIESKE, Rita; YU, Tiezheng; SU, Dan; XU, Yan; ISHII, Etsuko; BANG, Ye J.; MADOTTO, Andrea; FUNG, Pascale. Survey of Hallucination in Natural Language Generation. **ACM Computing Surveys**, v. 55, n. 12, p. 1-38, December 2023. Disponível em: <https://dl.acm.org/doi/abs/10.1145/3571730>. Acesso em: 18 abr. 2023.

KARUNANIDHI, Vignesh. How Long Did it Take ChatGPT to Reach 1 Million Users?. **Watcher Guru**, 29 jan. 2023. Business. Disponível em: <https://watcher.guru/news/how-long-did-it-take-chatgpt-to-reach-1-million-users>. Acesso em: 5 mar. 2023.

LATOUR, Bruno. **Ciência em ação**: como seguir cientistas e engenheiros sociedade afora. São Paulo: UNESP, 2000.

_____. **A Esperança de Pandora**: ensaios sobre a realidade dos estudos científicos. Bauru, SP: EDUSC, 2001.

_____. **Reagregando o Social**: uma introdução à teoria do Ator-Rede. Salvador. Edufba; São Paulo: EDUSC, 2012.

_____. **Investigação sobre os modos de existência**: uma antropologia dos modernos. Petrópolis,

RJ: Vozes, 2019.

LUGER, George F. **Inteligência Artificial**. Estruturas e Estratégias para a Solução de Problemas Complexos 4ª ed. Porto Alegre: Bookman, 2004.

MCKINSEY & COMPANY. What is generative AI?. In: **McKinsey Explainers**: Direct answers to complex questions. McKinsey Global Publishing, 19 jan. 2023. Disponível em: <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai>. Acesso em: 24 fev. 2023.

MIGALHAS. Advogado virtual? ChatGPT consegue "aprovação" na primeira fase da OAB. Migalhas, 22 fev. 2023. Disponível em: <https://www.migalhas.com.br/quentes/381875/advogado-virtual-chatgpt-consegue-aprovacao-na-primeira-fase-da-oab>. Acesso em: 7 mar. 2023.

NOLAN, Beatrice. Here are the schools and colleges that have banned the use of ChatGPT over plagiarism and misinformation fears. **Business Insider**, 30 jan. 2023. Disponível em: <https://www.businessinsider.com/chatgpt-schools-colleges-ban-plagiarism-misinformation-education-2023-1>. Acesso em: 5 mar. 2023.

RICH, Elaine; KNIGHT, Kevin. **Artificial Intelligence**. 2ª ed. Nova Iorque: McGraw-Hill, 1991.

SÉRVIO, Gabriel. ChatGPT escreveu centenas de livros vendidos na Amazon. **Olhar Digital**, 25 fev. 2023. Disponível em: <https://olhardigital.com.br/2023/02/25/reviews/chatgpt-escreveu-centenas-de-livros-vendidos-na-amazon/>. Acesso em: 7 mar. 2023.

SILVA, Mariana M., ChatGPT: inteligência artificial é aprovada em provas para médico, advogado e MBA nos EUA. **Exame**, 24 jan. 2023. Disponível em: <https://exame.com/future-of-money/chatgpt-inteligencia-artificial-e-aprovada-em-provas-para-medico-advogado-e-mba-nos-eua/>. Acesso em: 7 mar. 2023.

SYOZI, Ricardo. Desenvolvedores estão usando o ChatGPT para melhorar a história dos games. **Tecnoblog**, 3 fev. 2023. Disponível em: <https://tecnoblog.net/noticias/2023/02/03/desenvolvedores-estao-usando-o-chatgpt-para-melhorar-a-historia-dos-games/>. Acesso em: 7 mar. 2023.

_____. IA deixará que todos sejam programadores, diz CEO da Nvidia. **Tecnoblog**, 24 fev. 2023. Disponível em: <https://tecnoblog.net/noticias/2023/02/24/ia-deixara-que-todos-sejam-programadores-diz-ceo-da-nvidia/>. Acesso em: 7 mar. 2023.

VENTURA, Layse. ChatGPT passa em entrevista do Google e ganharia salário de quase R\$ 1

milhão por ano. **Olhar Digital**, 5 fev. 2023. Disponível em: <https://olhardigital.com.br/2023/02/05/pro/chatgpt-passa-em-entrevista-do-google-e-ganharia-salario-de-quase-r-1-milhao-por-ano/amp/>. Acesso em: 7 mar. 2023.

VINCENT, James. ChatGPT can't be credited as an author, says world's largest academic publisher. **The Verge**, 26 jan. 2023. Disponível em: <https://www.theverge.com/2023/1/26/23570967/chatgpt-author-scientific-papers-springer-nature-ban>. Acesso em: 7 mar. 2023.

WELSH, Matt. The End of Programming. **Communications of the ACM**, [s. l.], v. 66, n. 1, p. 34-35, Janeiro 2023. Disponível em: <https://cacm.acm.org/magazines/2023/1/267976-the-end-of-programming/>. Acesso em: 7 mar. 2023.

Recebido em 10/03/2023 – Aprovado em 02/05/2023