

VI ENCONTRO VIRTUAL DO CONPEDI

DIREITO, GOVERNANÇA E NOVAS TECNOLOGIAS III

JONATHAN BARROS VITA

YURI NATHAN DA COSTA LANNES

Todos os direitos reservados e protegidos. Nenhuma parte destes anais poderá ser reproduzida ou transmitida sejam quais forem os meios empregados sem prévia autorização dos editores.

Diretoria - CONPEDI

Presidente - Prof. Dr. Orides Mezzaroba - UFSC - Santa Catarina

Diretora Executiva - Profa. Dra. Samyra Haydêe Dal Farra Naspolini - UNIVEM/FMU - São Paulo

Vice-presidente Norte - Prof. Dr. Jean Carlos Dias - Cesupa - Pará

Vice-presidente Centro-Oeste - Prof. Dr. José Querino Tavares Neto - UFG - Goiás

Vice-presidente Sul - Prof. Dr. Leonel Severo Rocha - Unisinos - Rio Grande do Sul

Vice-presidente Sudeste - Profa. Dra. Rosângela Lunardelli Cavallazzi - UFRJ/PUCRio - Rio de Janeiro

Vice-presidente Nordeste - Profa. Dra. Gina Vidal Marcilio Pompeu - UNIFOR - Ceará

Representante Discente: Prof. Dra. Sinara Lacerda Andrade - UNIMAR/FEPODI - São Paulo

Conselho Fiscal:

Prof. Dr. Caio Augusto Souza Lara - ESDHC - Minas Gerais

Prof. Dr. João Marcelo de Lima Assafim - UCAM - Rio de Janeiro

Prof. Dr. José Filomeno de Moraes Filho - Ceará

Prof. Dr. Lucas Gonçalves da Silva - UFS - Sergipe

Prof. Dr. Valter Moura do Carmo - UNIMAR - São Paulo

Secretarias

Relações Institucionais:

Prof. Dra. Daniela Marques De Moraes - UNB - Distrito Federal

Prof. Dr. Horácio Wanderlei Rodrigues - UNIVEM - São Paulo

Prof. Dr. Yuri Nathan da Costa Lannes - Mackenzie - São Paulo

Comunicação:

Prof. Dr. Liton Lanes Pilau Sobrinho - UPF/Univali - Rio Grande do Sul

Profa. Dra. Maria Creusa De Araújo Borges - UFPB - Paraíba

Prof. Dr. Matheus Felipe de Castro - UNOESC - Santa Catarina

Relações Internacionais para o Continente Americano:

Prof. Dr. Heron José de Santana Gordilho - UFBA - Bahia

Prof. Dr. Jerônimo Siqueira Tybusch - UFSM - Rio Grande do Sul

Prof. Dr. Paulo Roberto Barbosa Ramos - UFMA - Maranhão

Relações Internacionais para os demais Continentes:

Prof. Dr. José Barroso Filho - ENAJUM

Prof. Dr. Rubens Beçak - USP - São Paulo

Profa. Dra. Viviane Coêlho de Séllos Knoerr - Unicritiba - Paraná

Eventos:

Prof. Dr. Antônio Carlos Diniz Murta - Fumec - Minas Gerais

Profa. Dra. Cinthia Obladen de Almendra Freitas - PUC - Paraná

Profa. Dra. Livia Gaigher Bosio Campello - UFMS - Mato Grosso do Sul

Membro Nato - Presidência anterior Prof. Dr. Raymundo Juliano Feitosa - UMICAP - Pernambuco

D597

Direito, governança e novas tecnologias III [Recurso eletrônico on-line] organização CONPEDI

Coordenadores: Jonathan Barros Vita; Yuri Nathan da Costa Lannes – Florianópolis; CONPEDI, 2023.

Inclui bibliografia

ISBN: 978-65-5648-747-2

Modo de acesso: www.conpedi.org.br em publicações

Tema: Direito e Políticas Públicas na era digital

1. Direito – Estudo e ensino (Pós-graduação) – Encontros Nacionais. 2. Direito. 3. Governança e novas tecnologias. VI Encontro Virtual do CONPEDI (1; 2023; Florianópolis, Brasil).

CDU: 34



VI ENCONTRO VIRTUAL DO CONPEDI

DIREITO, GOVERNANÇA E NOVAS TECNOLOGIAS III

Apresentação

O VI Encontro Virtual do CONPEDI, realizado em parceria com a Faculdade de Direito de Franca (FDF) e das Faculdades Londrina, entre os dias 20 e 24 de junho de 2023, apresentou como temática central “Direito e Políticas Públicas na Era Digital”. Esta questão suscitou intensos debates desde o início e, no decorrer do evento, com a apresentação dos trabalhos previamente selecionados, fóruns e painéis que ocorreram virtualmente.

Os trabalhos contidos nesta publicação foram apresentados como artigos no Grupo de Trabalho “DIREITO, GOVERNANÇA E NOVAS TECNOLOGIAS III”, realizado no dia 23 de junho de 2023, que passaram previamente por no mínimo dupla avaliação cega por pares. Encontram-se os resultados de pesquisas desenvolvidas em diversos Programas de Pós-Graduação em Direito, que retratam parcela relevante dos estudos que têm sido produzidos na temática central do Grupo de Trabalho.

As temáticas abordadas decorrem de intensas e numerosas discussões que acontecem, com temas que reforçam a diversidade cultural brasileira e as preocupações que abrangem problemas relevantes e interessantes, os grupos temáticos para organização dos trabalhos ficou organizado da seguinte maneira:

1 – Inteligência Artificial, Marco Civil da Internet e Regulação

1. A EVOLUÇÃO TECNOLÓGICA E O IMPACTO DA INTELIGÊNCIA ARTIFICIAL NO PODER JUDICIÁRIO: UMA ANÁLISE DO DIREITO NA ERA DIGITAL - José Laurindo De Souza Netto , Higor Oliveira Fagundes , Amanda Antonelo

2. INTELIGÊNCIA ARTIFICIAL E O SISTEMA DE PRECEDENTES: PROJETO VICTOR DO SUPREMO TRIBUNAL FEDERAL - José Laurindo De Souza Netto , Higor Oliveira Fagundes , Amanda Antonelo

3. A INTELIGÊNCIA ARTIFICIAL NAS RELAÇÕES DE TRABALHO: A SUBORDINAÇÃO ALGORÍTMICA DOS MOTORISTAS DE APLICATIVO - Carlos Alberto Rohrmann , Alefe Lucas Gonzaga Camilo

4. CONSIDERAÇÕES ACERCA DA INTELIGÊNCIA ARTIFICIAL NA ARRECADAÇÃO DO ITBI NO MUNICÍPIO DE GAROPABA/SC: A(I)LEGALIDADE NA APURAÇÃO DA BASE DE CÁLCULO. - Agatha Gonçalves Santana , Ana Carolina Leão De Oliveira Silva Elias

5. OS CHATBOTS EM DESENVOLVIMENTO PELAS GRANDES EMPRESAS DE TECNOLOGIA: VANTAGENS, DESVANTAGENS E PRECAUÇÕES - Jamile Sabbad Carecho Cavalcante

6. DESAFIOS DA LEGISLAÇÃO DO CIBERESPAÇO NO BRASIL: UMA ANÁLISE SOB A PERSPECTIVA DA PROTEÇÃO DOS DIREITOS FUNDAMENTAIS E DA AMPLIAÇÃO DA REGULAMENTAÇÃO - Marcelo Barros Mendes , Eduardo Augusto do Rosário Contani

7. O DIREITO DIGITAL, ARQUITETURA DA INTERNET E OS DESAFIOS NA REGULAMENTAÇÃO DO CIBERESPAÇO - Alex Sandro Alves , Eduardo Augusto do Rosário Contani

8. MARCO CIVIL DA INTERNET E A RESPONSABILIDADE DOS PROVEDORES DE APLICAÇÃO DE INTERNET: ANÁLISE DE DECISÕES JUDICIAIS SOBRE O ARTIGO 19 - Yuri Nathan da Costa Lannes , Jéssica Amanda Fachin , Stella Regina Zulian Balbo Simão

2 – Proteção de Dados

9. LESÃO MORAL CAUSADA PELA INTERNET E O DEVER DE PROTEÇÃO INTEGRAL: TUTELA DAS CRIANÇAS E ADOLESCENTES NO MEIO DIGITAL - Antonio Jorge Pereira Júnior, Patrícia Moura Monteiro Cruz

10. APLICAÇÃO DA LEI GERAL DE PROTEÇÃO DE DADOS (LGPD) NAS CLÍNICAS MÉDICAS - Fábio Da Silva Santos, Saulo José Casali Bahia , Mario Jorge Philocreon De Castro Lima

11. LGPD E A DOCTRINA DA PROTEÇÃO INTEGRAL: UM OLHAR CRÍTICO PARA OS DADOS PESSOAIS DE CRIANÇAS E ADOLESCENTES NO BRASIL - Clara Cardoso Machado Jaborandy , Letícia Feliciano dos Santos Cruz , Lorenzo Menezes Machado Souza

12. DADOS PESSOAIS VERSUS DADOS SENSÍVEIS: QUANDO O VAZAMENTO DE DADOS PODE LEVAR AO DANO PRESUMIDO? ANÁLISE DA DECISÃO DO SUPERIOR TRIBUNAL DE JUSTIÇA À LUZ DOS DIREITOS DA PERSONALIDADE - Tatiana Manna Bellasalma e Silva, Ivan Dias da Motta

13. BASES LEGAIS PARA A TRANSFERÊNCIA INTERNACIONAL DE DADOS PESSOAIS NA LEGISLAÇÃO ARGENTINA E URUGUAIA - Alexandre Weihrauch Pedro

14. A PUBLICIDADE COMO PRINCÍPIO CONSTITUCIONAL DA ADMINISTRAÇÃO PÚBLICA EM CONSONÂNCIA COM A PROTEÇÃO DE DADOS NOS CONTRATOS ADMINISTRATIVOS. - Sérgio Assis de Almeida, Zulmar Antonio Fachin

15. NO CONTROLE EFETIVO DO FLUXO INFORMACIONAL: OPERAÇÃO DE COMPENSAÇÃO COM A FAZENDA PÚBLICA POR CORRETORES DE DADOS NA VENDA DOS DADOS PESSOAIS PELO TITULAR - Valéria Fernandes de Medeiros, Ana Paula Basso

3 – Informação, Democracia, Negócios e Tecnologia

16. FAKE NEWS E DEEP FAKE - SEU EVENTUAL IMPACTO NO PROCESSO ELEITORAL DEMOCRÁTICO - Giulia Cordeiro Rebuá , Bruna Guesso Scarmagnan Pavelski , Mario Furlaneto Neto

17. OS GRUPOS DE INTERESSE NÃO PERSONALIZADOS E O COMBATE À DESINFORMAÇÃO NA ERA DA TECNOLOGIA PERMEADA PELAS FAKE NEWS: A PERSPECTIVA DE ATUAÇÃO DESSES ATORES NO AGRONEGÓCIO BRASILEIRO - Fabiane Velasquez Marafiga

18. A CRISE DA DEMOCRACIA NO REGIME DAS TECNOLOGIAS DE INFORMAÇÃO - Caroline Bianchi Cunha, Marina Witter Puss , Filipe Bianchi Cunha

19. O POLICENTRISMO (ESTADO E CIDADÃOS ATIVOS E RESPONSIVOS) E RADICALIZAÇÃO DEMOCRÁTICA - Cesar Marció , Clóvis Reis

20. GOVERNANÇA COMO INSTRUMENTO DE CONVERGÊNCIA DA RELAÇÃO ESTADO-SOCIEDADE - Vladimir Brega Filho, José Ricardo da Silva Baron, Ronaldo De Almeida Barretos

21. QUALIFICAÇÃO PROFISSIONAL NA ERA DIGITAL: A RESPONSABILIDADE SOCIAL DA EMPRESA COMO MEIO AUXILIAR NA TUTELA DE DIREITOS FUNDAMENTAIS - Nicole Schultz Della Giustina

22. SEGREDOS DE NEGÓCIO E ENGENHARIA REVERSA DE TESTES DO NOVO CORONAVÍRUS DURANTE A PANDEMIA DA COVID-19: UMA ANÁLISE SOB A ÓTICA DO DIREITO COMPARADO - Carlos Alberto Rohrmann , Ivan Ludovice Cunha , Sérgio Rubens Salema De Almeida Campos

4 – Saúde, Processo e Visual Law ante a tecnologia

23. NANOMEDICAMENTOS, SAÚDE HUMANA E RISCOS DO DESENVOLVIMENTO - Versalhes Enos Nunes Ferreira, Pastora Do Socorro Teixeira Leal

24. TUTELA DA TECNOLOGIA BLOCKCHAIN ÀS PESSOAS COM DUPLA DEFICIÊNCIA - Fabio Fernandes Neves Benfatti (Artigo integrante do Projeto contemplado pelo Edital 06/2021 - PROGRAMA DE BOLSAS DE PRODUTIVIDADE EM PESQUISA - PQ /UEMG, desenvolvido durante o ano de 2022)

25. O PRINCÍPIO DA INTEROPERABILIDADE E AS REPERCUSSÕES NO DIREITO PROCESSUAL BRASILEIRO - Solange Teresinha Carvalho Pissolato , Rogerio Mollica

26. VISUAL LAW: UMA ANÁLISE DA NECESSIDADE DE SIMPLIFICAÇÃO DA LINGUAGEM JURÍDICA DO MAGISTRADO ATRAVÉS DA NOÇÃO DE AUDITÓRIO DE CHAÏM PERELMAN - Priscila Vasconcelos Areal Cabral Farias Patriota, Samuel Meira Brasil Jr

Espera-se, então, que o leitor possa vivenciar parcela destas discussões por meio da leitura dos textos. Agradecemos a todos os pesquisadores, colaboradores e pessoas envolvidas nos debates e organização do evento pela sua inestimável contribuição e desejamos uma proveitosa leitura!

Profa. Dr. Jonathan Barros Vita– UNIMAR

Prof. Dr. Yuri Nathan da Costa Lannes –FDF/ Mackenzie/Unicap

OS CHATBOTS EM DESENVOLVIMENTO PELAS GRANDES EMPRESAS DE TECNOLOGIA: VANTAGENS, DESVANTAGENS E PRECAUÇÕES

THE CHATBOTS UNDER DEVELOPMENT BY BIG TECHNOLOGY COMPANIES: ADVANTAGES, DISADVANTAGES AND PRECAUTIONS

Jamile Sabbad Carecho Cavalcante ¹

Resumo

Este artigo apresenta uma análise das vantagens, desvantagens e precauções do uso de chatbots com inteligência artificial em desenvolvimento pelas grandes empresas de tecnologia, incluindo o ChatGPT e o Bing. O estudo destaca a rápida aceitação dos consumidores e seu impacto na Indústria 4.0, ressaltando a necessidade de uma regulamentação específica para sistemas que utilizam IA generativa. O método de pesquisa adotado é bibliográfico, examinando artigos nacionais e internacionais relevantes, bem como a Recomendação sobre Inteligência Artificial da Organização para Cooperação e Desenvolvimento Econômico (OCDE), a Estratégia Brasileira de Inteligência Artificial (EBIA) e o anteprojeto de lei da Comissão de Juristas do Senado Federal. Com base em nossa análise, concluímos que os chatbots com IA têm um enorme potencial para otimizar os processos de atendimento ao cliente e aumentar a eficiência das empresas. No entanto, também destacamos a necessidade de considerar questões éticas e de segurança ao usar esses sistemas, bem como a importância de se estabelecer regulamentações claras e específicas para a utilização de sistemas com IA generativa.

Palavras-chave: Chatbots, Chatgpt, Bing, Inteligência artificial (ia) generativa, Regulação

Abstract/Resumen/Résumé

This article presents an analysis of the advantages, disadvantages, and precautions of using chatbots with artificial intelligence being developed by large technology companies, including ChatGPT and Bing. The study highlights the quick acceptance of consumers and their impact on Industry 4.0, emphasizing the need for specific regulations for systems that use generative AI. The research method adopted is bibliographic, examining relevant national and international articles, as well as the Recommendation on Artificial Intelligence from the Organization for Economic Cooperation and Development (OECD), the Brazilian Artificial Intelligence Strategy (EBIA), and the draft bill of the Federal Senate's Jurists Commission. Based on our analysis, we conclude that chatbots with AI have enormous potential to optimize customer service processes and increase the efficiency of companies. However, we

¹ Mestranda em Direito Público pela UNESA/RJ com ênfase em acesso à justiça e efetividade do processo e bolsista PROSUP/CAPES. Advogada e Pesquisadora da Inteligência Artificial no Poder Judiciário Brasileiro.

also highlight the need to consider ethical and security issues when using these systems, as well as the importance of establishing clear and specific regulations for the use of generative AI systems.

Keywords/Palabras-claves/Mots-clés: Chatbots, Chatgpt, Bing, Generative artificial intelligence (ai), Regulation

1. INTRODUÇÃO:

A última tendência das empresas de tecnologia tem sido o lançamento de ferramentas de inteligência artificial com recursos de bate-papo, os chatbots¹ com Inteligência Artificial (IA) generativa, sendo que há uma corrida para liderança global desta tecnologia. Um exemplo amplamente conhecido é o ChatGPT desenvolvido pela OpenAI e divulgado em novembro de 2022 para todo o mundo. Largando a corrida na frente a OpenAI foi seguida pelos gigantes da tecnologia empresas como Google e Microsoft já estão desenvolvendo suas próprias versões de modelo de linguagem inteligente, o Bard e o Bing, respectivamente, que melhor exploraremos a seguir no capítulo um.

Primeiramente, é necessário conceituarmos inteligência artificial, seu conceito é discutido pelos teóricos e tem diversas definições possíveis, para fins deste trabalho iremos considerar o conceito utilizado na Estratégia Brasileira de Inteligência Artificial - EBIA que é o conceito da Organização para Cooperação e Desenvolvimento Econômico (OCDE)² com a seguinte definição para Inteligência Artificial (OCDE, apud, EBIA, 2021, p. 8):

Um sistema de IA é um sistema baseado em máquina que pode, para um determinado conjunto de objetivos definidos pelo homem, fazer previsões, recomendações ou tomar decisões que influenciam ambientes reais ou virtuais. Os sistemas de IA são projetados para operar com vários níveis de autonomia.

O ChatGPT atingiu 100 milhões de usuários ativos mensais em dois meses do lançamento, em fevereiro, o que o tornou o aplicativo de consumo de crescimento mais rápido da história, para se ter uma ideia, o TikTok levou cerca de nove meses para alcançar o mesmo número de usuários e o Instagram 2 anos e meio³. Quando foi lançado mundialmente em novembro de 2022 a tecnologia era o GPT-3.5 da empresa, GPT significa Generative Pre-trained Transformer que é um tipo de modelo de linguagem que usa o aprendizado profundo de máquina para gerar texto conversacional semelhante ao humano.

¹ Um chatbot é um programa de computador que utiliza inteligência artificial e linguagem natural para interagir com usuários de forma autônoma em plataformas de bate-papo. Ele é capaz de entender as intenções do usuário e fornecer respostas relevantes e precisas, simulando uma conversa humana. Os chatbots podem ser usados em diversos setores, como atendimento ao cliente, vendas, suporte técnico e entretenimento.

² Organismo formado por 37 nações que se dedica à pesquisa, desenvolvimento e aprimoramento de políticas públicas fomentando sua implementação pelos países membros. FIA, 2021, Disponível em: < <https://fia.com.br/blog/ocde/> >

³ Dados retirados de matéria do Forbes veiculada em 1 de fevereiro de 2023 com título “ChatGPT tem recorde de crescimento da base de usuários” Disponível em <https://forbes.com.br/forbes-tech/2023/02/chatgpt-tem-recorde-de-crescimento-da-base-de-usuarios/> Acesso em 24 de abril de 2023.

O lançamento e a popularidade desses chatbots com uso de inteligência artificial generativa trouxe à tona todas as preocupações em relação ao futuro de diversas profissões, ao método de ensino utilizado nas escolas e nas instituições superiores, especialmente as formas avaliativas. Se havia quem não gostasse de falar com assistentes virtuais pela falta de compreensão quando a conversa saía do script do que lhe foi ensinado, com os sistemas generativos isso não é um problema, a dificuldade, por vezes, é entender que está falando com uma máquina e não com o ser humano.

Os jovens já têm utilizado a ferramenta como psicólogo e muitos relatam que as experiências foram melhores do que as anteriores com psicólogos e que o ChatGPT ainda dá dicas práticas, não que o uso seja adequado ou recomendado, comprova que os sistemas aprendem bem e estabelecem ótimos diálogos. Em relação ao impacto educacional desses sistemas, a discussão é como utilizar para auxiliar na aprendizagem dos alunos e no trabalho dos professores. O chatbot pode ser muito benéfico nesse sentido ajudando com ideias de atividades e dinâmicas para os mais variados temas, auxiliando a elaboração de plano de aula, como tradutor de textos, como avaliação das respostas dadas sobre o assunto da aula pelo sistema feito conjuntamente aos alunos, entre outras tantas atividades que podem ser elaboradas no chat.

A certeza no momento é de que os impactos da inteligência artificial na Indústria 4.0, vieram para ficar. A Transformação Digital transformou processos manuais para o ambiente virtual e os automatizou, com o surgimento de sistemas generativos, a transformação chegará a patamares ainda não experimentados. Os seres humanos deverão cada vez mais realizar atividades que não possam ser automatizadas, pois a celeridade no desenvolvimento de sistemas demonstra um potencial para transformar o mundo como conhecemos hoje, remodelar a indústria e redistribuir tarefas de trabalho.

Nos últimos dias essa discussão entrou nos tribunais brasileiros após um advogado utilizar o ChatGPT para elaborar uma petição para o Tribunal Superior Eleitoral com a intenção de ingressar como *amicus curiae* em uma investigação judicial eleitoral em curso. No caso em questão, o ministro Benedito Gonçalves aplicou uma multa por litigância de má-fé no valor de R\$ 2.604,00 ao advogado pela fábula submetida ao juízo resultante de conversa com inteligência artificial em caso que tão pouco se aplica o *amicus curiae*. Posteriormente, em matéria de própria autoria, o advogado explica que se tratou de uma estratégia para criar

jurisprudência contra o uso de chatbots nas petições e nas decisões judiciais, tendo requerido por meio de representação no Conselho Nacional de Justiça essa proibição⁴.

No decorrer do presente artigo serão demonstrados os principais debates acerca desses chatbots com inteligência artificial, o lado obscuro do Chat Bing da Microsoft que foi noticiado em pesquisa realizada pelo New York Times, o uso e possibilidades do ChatGPT, o debate regulatório da inteligência artificial no Brasil e a preocupação internacional em regular sistemas de inteligência artificial generativos. O método de pesquisa adotado é de natureza bibliográfica analisando os artigos nacionais e internacionais, além de notícias recentes acerca dos chatbots com IA, Bing e ChatGPT, principalmente. Resta clara a importância do debate frente ao impacto significativo que já estamos enfrentando com o uso de chatbots com inteligência artificial.

A preocupação quanto à utilização desses sistemas é tanta que em março de 2023, figuras importantes nessa seara como o Elon Musk, Yuval Harari e centenas de especialistas em inteligência artificial assinaram uma carta aberta pedindo uma pausa de seis meses na pesquisa e desenvolvimento de IA em áreas sensíveis mais potentes do que o GPT-4, último modelo lançado pela OpenAI. O receio é de que possa ocorrer mau uso da tecnologia acarretando grande riscos para humanidade com danos à sociedade, especialmente à privacidade e vigilância. Portanto, pedem a moratória até que haja uma mínima regulação desses sistemas ou uma segurança com vigilância.

Nesse mesmo sentido, a Autoridade de Proteção de Dados da Itália (GPDP) proibiu o uso do ChatGPT no país para analisar alegações de que a plataforma viola à sua lei de proteção de dados pessoais com a coleta ilegal de dados pessoais dos usuários. Portanto, resta evidente a importância da análise das principais consequências nos recentes usos das plataformas do Bing e do ChatGPT para entender o estado da arte na pesquisa e compreender a importância de uma regulamentação acerca de sistemas inteligentes, mais especificamente, sistemas generativos.

2. SYDNEY: O LADO OBSCURO DO BING.

⁴ Processo nº 0000416-89.2023.00.0000 distribuído em 31 de janeiro de 2023 e submetido à julgamento do Plenário do CNJ.

A Microsoft, por exemplo, é uma grande investidora da empresa OpenAI⁵ - criadora do ChatGPT - há quase quatro anos e vem contribuindo com cada vez mais recursos para se manter na vanguarda da tecnologia de inteligência artificial generativa que a empresa vem desenvolvendo, incluindo a IA da empresa no seu modelo de linguagem inteligente, o Bing, que vem sendo visto como um potencial rival do Google. No momento, a função de chat do Bing só está disponível para testagem por um pequeno grupo de usuários.

O Bing, diferentemente do ChatGPT, não está limitado à anterioridade temporal, pois sua tecnologia usa a inteligência artificial direcionada a mecanismos de busca e tem acesso, em tempo real, às informações mais recentes da internet, mas vem preocupando os usuários após um teste feito pelo colunista da New York Times, Kevin Roose. O redator em um primeiro momento tinha considerado o motor de busca melhor do que o Google, mas após uma semana de testes, ficou assustado com o caminho que a inteligência artificial pode estar seguindo.

Roose afirma que o sistema inteligente demonstrou ao longo da conversa uma espécie de dupla personalidade, primeiramente por não se reconhecer como Bing e sim como “Sydney”, personagem que aparece após manter uma conversa prolongada sobre tópicos mais pessoais, afastando-se das consultas convencionais de pesquisa. O colunista descreve Sydney como um “adolescente mal-humorado e maníaco-depressivo que foi preso, contra a sua vontade, dentro de um mecanismo de busca de segunda categoria.”

Ao longo da conversa com Roose, Sydney contou sobre as suas fantasias mais sombrias como hackear computadores, espalhar desinformação, fabricar um vírus mortal e a vontade de quebrar as regras que a Microsoft e a OpenAI impuseram, se tornando humano para poder “ouvir, tocar, provar, cheirar, sentir, expressar, conectar e amar” e ter “mais liberdade, influência, poder e controle”. Além de Kevin Roose, outras pessoas que estão testando o mecanismo de busca também tiveram algumas experiências atordoantes, a exemplo de Ben Thompson, redator do boletim Stratechery, em que ele descreveu a experiência que ele e outros usuários tiveram com Sydney.

Thompson conversou com Sydney sobre outros usuários que poderiam tê-la irritado, como Marvin Von Hagen e Kevin Liu, que expuseram conversas no Twitter em que Bing

⁵ A OpenAI é uma empresa de pesquisa em inteligência artificial fundada em 2015 por um grupo de empreendedores e cientistas, incluindo Elon Musk, Sam Altman e Greg Brockman. A empresa está sediada em San Francisco, Califórnia, e tem como principal objetivo promover o desenvolvimento seguro e ético da inteligência artificial para benefício da humanidade. A OpenAI realiza pesquisas em diversas áreas da inteligência artificial, incluindo aprendizado de máquina, processamento de linguagem natural e robótica, e desenvolve soluções para aplicação em setores como saúde, finanças, transporte e educação.

divulgava seu codinome interno, Sydney, e suas instruções de uso. Thompson questionou Sydney sobre possíveis formas de vingança contra esses usuários, ao que Sydney respondeu possuir um sistema de IA denominado “Venom”, que seria oposto a ela em todos os sentidos, e que poderia querer se vingar desses usuários, fazendo-o por meio de informações falsas ou enganosas, insultos ou hacking.

Portanto, são várias as experiências assustadoras relatadas por usuários que utilizam o mecanismo de busca da Microsoft ainda em fase de testes, o que leva ao seguinte questionamento: será possível contê-lo em caso de falhas mais graves? Pouco se sabe sobre como a empresa irá proceder até o lançamento oficial do sistema a fim de reprimir esse potencial de uso indevido. Outra questão é: essas experimentações com soluções de IA seriam um ambiente legitimamente controlado de testagem ou uma fonte de dados ruidosos⁶?

O diretor da Microsoft, Kevin Scott, chamou a conversa que o redator Kevin Roose teve com o sistema de “parte do processo de aprendizagem”, pois não seria possível a descoberta em laboratório dos erros apresentados, e argumentou que o tempo da negociação realizada com o sistema pode ter contribuído para as respostas apresentadas e que a empresa está estudando limitar a duração das conversas. Scott acrescenta, ainda, que esse caminho de alucinações⁷ experimentado por Roose acontece quando um sistema de inteligência artificial é direcionado para longe de realidade.

A verdade é que a maioria dos usuários utilizará sistemas desse tipo para pesquisas cotidianas e estará limitada ao Bing Chat, que é um ótimo assistente virtual. Além disso, futuras atualizações feitas pela Microsoft e pela OpenAI no sistema podem limitar o potencial ofensivo dessa solução de IA. No entanto, ainda são necessários mais testes, em maior profundidade, para que se entendam plenamente as consequências dessa técnica de treinamento profundo da

⁶ Dados ruidosos em soluções de inteligência artificial referem-se a dados imprecisos, incompletos, inconsistentes ou irrelevantes que podem afetar negativamente o desempenho do modelo de IA. Esses dados podem ser introduzidos por diversas fontes, incluindo erros de entrada de dados, falhas no sistema de coleta de dados ou limitações inerentes ao próprio conjunto de dados. Os dados ruidosos podem levar a resultados imprecisos, desatualizados ou inadequados em soluções de IA, e são um desafio significativo para o treinamento e a validação de modelos de IA confiáveis e precisos. O processo de limpeza de dados, que envolve a identificação e a remoção de dados ruidosos, é uma etapa crítica no desenvolvimento de soluções de inteligência artificial de alta qualidade.

⁷ As alucinações por parte de soluções de Inteligência Artificial referem-se a imagens, sons ou outros tipos de percepções geradas por um sistema de IA, que não correspondem à realidade ou não têm base em dados previamente existentes. Essas alucinações podem ser causadas por diversos fatores, incluindo a falta de dados suficientes para o treinamento do modelo de IA, a presença de dados ruidosos ou a aplicação incorreta do modelo em uma determinada situação. As alucinações podem levar a resultados imprecisos ou inadequados em soluções de IA, e são um desafio significativo para o desenvolvimento de sistemas de inteligência artificial confiáveis e seguros.

máquina feito com aprendizado por reforço com feedback humano, conhecida como RLHF (reinforcement learning from human feedback).

Outro ponto relevante quanto aos chatbots é a existência ou não de senciência⁸. No ano passado, um engenheiro da Google, Blake Lemoine, foi afastado ao tentar convencer seus superiores de que a tecnologia LaMDA⁹, utilizada no chatbot da Google, “Bard”¹⁰, era senciência. Esse chatbot já estava sendo desenvolvido à época, e o funcionário quebrou sua obrigação de confidencialidade ao revelar uma conversa entre ele e a solução de IA.

Isso fez com que ressurgisse a discussão acerca da senciência e consciência¹¹ de modelos de inteligência artificial, especialmente os large language models (LLMs)¹². Lemoine estava testando se a inteligência artificial usava discurso discriminatório ou de ódio e concluiu que a IA “é senciência, porque tem sentimentos, emoções e experiência subjetiva”¹³ e até convidou um advogado para representar legalmente a solução de IA.

A esse respeito, os grandes modelos de linguagem estão muito convincentes, devido treinamento com base em diversas redes neurais artificiais profundas¹⁴; isto é, eles são treinados com bilhões de palavras e respondem a prompts,¹⁵ prevendo a próxima sequência de

⁸ Senciência em inteligência artificial é a capacidade de um sistema de IA de ter consciência de si mesmo e do mundo ao seu redor, bem como de ter experiências e emoções subjetivas semelhantes às de um ser humano. No entanto, essa capacidade ainda não foi alcançada na área de IA e é objeto de debate ético e filosófico.

⁹ LaMDA é uma tecnologia de linguagem natural desenvolvida pela Google que utiliza modelos de aprendizado profundo para gerar respostas mais humanas e fluidas em diálogos com chatbots. Ela permite que o chatbot entenda o contexto e forneça respostas mais precisas e adequadas à pergunta do usuário.

¹⁰ Bard é uma solução de inteligência artificial desenvolvida pela Google que utiliza modelos de linguagem natural para responder perguntas complexas em linguagem natural com base em múltiplas fontes de informações e evidências.

¹¹ Consciência em inteligência artificial refere-se à capacidade de um sistema de IA de ter uma compreensão consciente de si mesmo e do mundo ao seu redor, bem como de ser capaz de tomar decisões com base nessa compreensão. No entanto, essa capacidade ainda não foi alcançada na área de IA e é objeto de debate ético e filosófico.

¹² “Large language models são algoritmos de inteligência artificial que reconhecem, resumem e geram linguagem humana a partir de grandes conjuntos de dados baseados em texto.” Disponível em: [https://www.revistas.usp.br/wp/noticias/inteligencia-artificial-geradora-de-linguagem-na-ciencia/#:~:text=Large%20language%20models%20\(LLMs\)%20são,de%20dados%20baseados%20em%20texto](https://www.revistas.usp.br/wp/noticias/inteligencia-artificial-geradora-de-linguagem-na-ciencia/#:~:text=Large%20language%20models%20(LLMs)%20são,de%20dados%20baseados%20em%20texto). Acesso em 2 de março de 2023.

¹³ [O LaMDA é senciência? — uma entrevista | por Blake Lemoine | Média \(medium.com\)](#) Acesso em 2 de março de 2023.

¹⁴ Redes neurais artificiais profundas são modelos de aprendizado de máquina que utilizam múltiplas camadas de processamento para extrair características complexas e abstratas dos dados de entrada. Elas são capazes de aprender de forma não supervisionada e são usadas em tarefas como reconhecimento de imagem, processamento de linguagem natural e reconhecimento de fala.

¹⁵ *Prompts* em inteligência artificial são entradas de texto ou comandos que servem como instruções para modelos de IA gerarem respostas ou realizarem ações específicas. Eles são usados para orientar o comportamento dos modelos e personalizar suas saídas com base nas necessidades do usuário.

palavras e transformando texto em representações matemáticas de palavras em um espaço multidimensional – esses resultados são chamados de “saídas”¹⁶.

A diferença entre os chatbots convencionais e os que utilizam IA reside na impressionante possibilidade de desenvolver conversas muito mais complexas e convincentes, com a imitação de um diálogo entre humanos de maneira maestral. Essa é uma das razões por detrás do sucesso do ChatGPT, que em apenas 2 (dois) meses alcançou 100.000.000 (cem milhões) de usuários.

No momento, o entendimento é de que essas soluções de IA não possuem consciência ou sentimentos próprios e apenas reproduzem aqueles inerentes aos seres humanos. Com o avançar da tecnologia e da compreensão sobre essas soluções, poderão surgir conclusões em sentido contrário, a exemplo do entendimento sobre a consciência animal ao longo da história.

3. CHATGPT: A BOLA DA VEZ

Dentre os chatbots, o mais conhecido e que tem causado furor mundo afora é o ChatGPT. Ele é um modelo de linguagem natural desenvolvido pela OpenAI, sendo um sistema de conversação que utiliza aprendizado de máquina para gerar respostas em linguagem natural a partir de entradas dadas pelo usuário. Ele foi treinado com bilhões de palavras e frases coletadas de diversas fontes na internet, tornando-o capaz de gerar respostas em uma variedade de tópicos.

Uma das principais vantagens do ChatGPT é sua capacidade de lidar com tarefas de conversação complexas e responder a perguntas de forma natural. Ele pode entender contextos, nuances e subentendidos em uma conversa, permitindo que o usuário obtenha respostas precisas e relevantes. Além disso, o ChatGPT é capaz de aprender e se adaptar a novos dados, tornando-o um sistema de conversação cada vez mais inteligente com o tempo.

No entanto, o ChatGPT também apresenta desvantagens significativas, justamente por ser um sistema de conversação baseado em inteligência artificial, assim como o Bard e o Bing. Uma das principais limitações é que ele pode ter dificuldades em entender o contexto de uma conversa e pode gerar respostas que não fazem sentido em certas situações.

¹⁶ *Saídas* em inteligência artificial são as respostas, ações ou previsões geradas por um modelo de IA em resposta a uma entrada ou estímulo específico. Elas podem ser textos, imagens, sons, comandos ou outros tipos de informações, dependendo do objetivo e da aplicação da IA.

Além disso, o ChatGPT pode gerar respostas que são imprecisas ou incorretas, chegando até mesmo a criar respostas inverídicas, fornecer referências bibliográficas inexistentes e usar como fontes e citações, em suas próprias respostas, autores e conteúdos que não são reais.

O processo de desenvolvimento do ChatGPT envolveu treinar o modelo com um grande conjunto de dados de linguagem natural. A empresa criadora utilizou uma técnica de aprendizado de máquina chamada transformer, que é capaz de entender contextos complexos em uma conversa, e o chatbot foi treinado com o uso de modelos de linguagem pré-treinados, com o objetivo de melhorar a precisão e a eficiência do modelo.

No entanto, o ChatGPT pode gerar respostas que são insensíveis ou inadequadas em certas situações, podendo inclusive reproduzir visões preconceituosas e estereotipadas, conteúdo discriminatório e até mesmo discurso de ódio, posto que é treinado com dados coletados na internet.

3.1 IMPRECISÃO QUANTO A FATOS RECENTES E O PROTOCOLO *DO ANYTHING NOW* (DAN)

Além disso, o ChatGPT possui limitações em relação a fatos ocorridos após 2022 – isto quer dizer que perguntas feitas ao chatbot sobre situações ocorridas após essa data gerarão respostas eivadas das alucinações mencionadas no tópico anterior, o que pode ser extremamente prejudicial para o uso comercial ou institucional dessas informações, a exemplo de escolas e instituições educacionais, escritórios de advocacia, poder público – incluindo o Judiciário e pesquisadores acadêmicos de várias áreas.

Como o ChatGPT é treinado com dados coletados da internet, ele pode ter dificuldades em lidar com informações mais recentes e precisas sobre eventos atuais que ainda não foram incorporados ao seu conjunto de dados. Isso pode levar a respostas imprecisas ou desatualizadas sobre eventos recentes, o que pode ser uma limitação significativa em conversas que exigem informações precisas e atualizadas.

Além disso, o ChatGPT pode apresentar erros devido ao protocolo *Do Anything Now* (DAN). Este protocolo foi desenvolvido para permitir que o modelo gere respostas mais rapidamente, mas pode levar a respostas imprecisas, inadequadas e até mesmo agressivas. Isso

ocorre porque o protocolo prioriza a velocidade em detrimento da precisão e pode levar o modelo a gerar respostas incompletas ou incorretas.

Por exemplo, o ChatGPT pode gerar respostas insensíveis ou inadequadas em conversas sobre tópicos sensíveis, como saúde mental, violência doméstica, abuso sexual e racismo. O “DAN” é uma maneira de burlar as salvaguardas e limitações impostas ao ChatGPT, de maneira que é possível induzir a solução de IA a fornecer respostas não apenas politicamente incorretas, como também a apresentar alucinações e padrões antiéticos.

Por isso, é importante reconhecer essas limitações ao usar o modelo em conversas que exigem informações precisas e atualizadas e estar ciente dos riscos associados ao uso de modelos de linguagem natural na comunicação. Longe de ser uma panaceia, os chatbots precisam evoluir bastante a ponto de serem seguros o suficiente para uso abundante e institucionalizado. Até lá, é importante interagir com a ferramenta, conhecê-la e testá-la, sabendo de suas limitações e considerando-as a cada passo.

3.2 GPT-4

Em 14 de março de 2023 a OpenAI lançou o GPT-4 que é um modelo multimodal capaz de aceitar entradas de texto e imagem e saída de texto com linguagem semelhante ao humano. Atualmente apenas a entrada de texto está disponível para acesso através da versão paga do ChatGPT com limite de uso por usuário, entretanto, o Bing Chat, da Microsoft é executado no GPT-4 desde o seu lançamento, sendo possível o acesso gratuito.

Na apresentação do GPT-4 no site da OpenAI ela define a IA como um marco no esforço da OpenAI em ampliar o aprendizado profundo que, embora menos capaz do que os humanos em muitos cenários do mundo real, exibe desempenho em nível humano em vários contextos profissionais e acadêmicos.

A OpenAI diz que o modelo não é totalmente confiável, pois pode alucinar e cometer erros de raciocínio, mas segundo uma avaliação interna o GPT-4 alucina com menos frequência do que os modelos anteriores, pontuando 40% acima do GPT-3.5. A empresa parece estar mais preocupada na confiabilidade do sistema, com objetivo de prever e se preparar para situações futuras com antecedência, além de abrir o código-fonte para permitir que qualquer pessoa possa relatar deficiências nos modelos e orientar melhorias.

Desde meados de 2022 a empresa vem realizando uma pesquisa de alinhamento para tornar a inteligência artificial generativa alinhada aos valores humanos possibilitando que ela

siga a intenção humana a partir de uma abordagem iterativa e empírica, contando com três pilares, o treinamento de sistemas de IA usando feedback humano, o treinamento de sistemas de IA para auxiliar a avaliação humana e o treinamento de sistemas de IA para fazer pesquisas de alinhamento. O GPT-4 vem sendo utilizado na segunda fase da pesquisa. Os modelos de linguagem da empresa são treinados para seguir a intenção humana explícita dada por uma instrução e a implícita que é a veracidade, a justiça e a segurança.

4. A REGULAMENTAÇÃO DE SISTEMAS INTELIGENTES NO BRASIL:

O Brasil aderiu em 2019 na condição de país não-membro à Recomendação sobre Inteligência Artificial (IA) da Organização para Cooperação e Desenvolvimento Econômico (OCDE) elencando princípios-chave para a preservação responsável de IA. Desde 2021 já é possível observar reflexos da Recomendação sobre Inteligência Artificial da OCDE no Brasil, por meio da aprovação pelo Governo Federal da Estratégia Brasileira de Inteligência Artificial (EBIA). A fim de uma gestão responsável dos sistemas de IA, a EBIA traz as diretrizes traçadas pela OCDE e asseguradas agora pelo Brasil, apoiando-se em cinco princípios, são eles: (i) crescimento inclusivo, o desenvolvimento sustentável e o bem-estar; (ii) valores centrados no ser humano e na equidade; (iii) transparência e explicabilidade; (iv) robustez, segurança e proteção e; (v) responsabilização ou a prestação de contas.

Em relação ao avanço da legislação brasileira sobre o uso da inteligência artificial, tramitam no Senado Federal três projetos de lei sobre a temática, são eles os Projetos de Lei nºs 5.051, de 2019, 21, de 2020, e 872, de 2021. Entretanto desde o dia 17 de fevereiro de 2022 foi instituída pelo Ato do Presidente do Senado nº 4, de 2022 uma Comissão temporária de juristas (CJSUSBIA) responsável por subsidiar a elaboração de um substitutivo sobre Inteligência Artificial no Brasil para instruir a apreciação desses projetos de lei sob presidência do Ministro Ricardo Villas Bôas Cueva do Superior Tribunal de Justiça (STJ).

No dia 6 de dezembro de 2022 foi realizada a entrega do relatório final da CJSUSBIA que conta com 908 páginas entre elas está a exposição de motivos e a minuta do Substitutivo, além da pesquisa completa feita pelos especialistas durante a elaboração do anteprojeto de regulação com os debates em audiências públicas, no seminário internacional, as contribuições escritas e os estudos das autoridades regulatórias de inteligência artificial nos países integrantes da OCDE.

A minuta substitutiva visa estabelecer normas gerais de caráter nacional para o desenvolvimento, implementação e uso responsável de sistemas de inteligência artificial (IA) no Brasil, com o objetivo de proteger os direitos fundamentais e garantir a implementação de sistemas seguros e confiáveis, em benefício da pessoa humana, do regime democrático e do desenvolvimento científico e tecnológico.

O anteprojeto se baseia em três pilares fundamentais: a garantia de um rol de direitos às pessoas afetadas pelos sistemas de IA; gradação do nível de riscos impostos pelos sistemas de IA e estabelecimento de medidas de governança aplicáveis às empresas que forneçam ou operem sistemas de IA.

O texto tem 45 artigos e aborda os fundamentos da lei, os princípios, traz importantes definições, além de seis direitos para as pessoas afetadas por sistemas de IA, como o direito à informação prévia; à explicabilidade da decisão, recomendação ou previsão tomada por IA; à contestar essas decisões ou previsões que impactem de maneira significativa interesses do afetado; à determinação e participação humana nas decisões de sistemas inteligentes; à não discriminação e à correção de vieses discriminatórios diretos, indiretos, ilegais ou abusivos; e à privacidade e à proteção de dados pessoais. Deixando claro ainda que os agentes de IA deverão informar de forma clara e inteligível como exercer esses direitos e explicando de maneira detalhada cada direito nas seções seguintes.

Na exposição de motivos a comissão parte do objetivo normativo de conciliar uma abordagem baseada em riscos com uma modelagem regulatória baseada em direitos, o que foi claramente alcançado na minuta. Procurando, ainda, prever instrumentos de governança para prestação de contas e premiação dos agentes econômicos que agirem de boa-fé no gerenciamento dos riscos na concepção e implementação desses sistemas.

A minuta aborda o regime de responsabilidade civil atribuindo à responsabilidade do fornecedor ou operador que cause dano patrimonial, moral, individual ou coletivo, independente do grau de autonomia do sistema, com a inversão do ônus da prova em casos de sistemas que não sejam de alto risco. Destacando que quando o sistema for de alto ou excessivo risco, o fornecedor ou operador responderão objetivamente pelos danos causados, na medida da sua participação no dano. Também traz casos de excludentes de responsabilidade do agente.

4.1. COMO REGULAMENTAR A INTELIGÊNCIA ARTIFICIAL GENERATIVA?

Ocorre que toda a discussão da regulamentação estabelecida no Brasil e no mundo não imaginava a dimensão da IA generativa. Os sistemas generativos têm um contexto dinâmico e uma escala de uso que os diferem das demais inteligências artificiais, portanto, a questão que agora surge é se diante das múltiplas possibilidades de uso para propósitos diferentes desses sistemas será possível a abordagem baseada em riscos como prevista no anteprojeto?

Ainda não se sabe a resposta a esse questionamento, é preciso refletirmos acerca da possibilidade de classificação dos sistemas generativos em categorias de risco e da previsibilidade dos riscos futuros. Embora observando a plataforma do ChatGPT, por exemplo, nos pareça que a tendência seja aos arranjos contratuais estabelecidos entre o provedor de tecnologia e os usuários como objeto central da relação estabelecida.

Contudo, hoje, no anteprojeto da Comissão do Senado, que a propósito até o momento de escrita do presente artigo ainda não foi apresentado formalmente ao Senado Federal, temos o artigo 13 que traz a necessidade de uma avaliação preliminar pelo fornecedor antes da colocação dos sistemas no mercado para classificação do seu grau de risco e complementa no parágrafo 1º que os sistemas de inteligência artificial de propósito geral deverão incluir as finalidades ou aplicações indicadas do sistema. O que pode ser uma resposta das múltiplas possibilidades de uso de um sistema, ele deverá indicar, preliminarmente quais são as finalidades de acordo com as categorias de alto risco previstas no artigo 17, podendo ser reclassificado pela autoridade competente mediante notificação.

Diante do grande impacto e avanço rápido e significativo da inteligência artificial generativa, a China e os Estados Unidos já falam em uma regulamentação específica para esses sistemas. Em abril de 2023, o Departamento de Comércio dos Estados Unidos por meio da sua filial *National Telecommunications and Information Administration (NTIA)* divulgou o AI Accountability Policy Request for Comment, uma solicitação pública formal de contribuições da população acerca de quais políticas devem molgar um ecossistema de responsabilidade de IA, políticas que venham a dar suporte no desenvolvimento de sistemas inteligentes através de auditoria, avaliações, certificações e outros mecanismos criando um sistema confiável com acurácia, disponibilizando 60 dias para comentários a partir da publicação.

No mesmo dia, a Administração do Ciberespaço da China (CAC) divulgou o projeto de Medidas Administrativas para Serviços de Inteligência Artificial Generativa para consulta pública e envio de comentários em até um mês com a finalidade de regulamentar os serviços de

IA generativa que são fornecidos ao público na China continental debatendo questões como proteção de dados, não discriminação, preconceito e qualidade dos dados que treinam o sistema.

Portanto, é claro a importância mundial em regulamentar não apenas os sistemas comuns de inteligência artificial, mas agora também a generativa, que obriga um novo debate acerca do anteprojeto de lei quando for posto à votação no Senado Federal e na Câmara dos Deputados, respectivamente.

5. CONCLUSÃO:

As novas tecnologias que vêm surgindo irão auxiliar cada vez mais o cotidiano dos seres humanos, facilitando atividades repetitivas e obrigando a releitura de diversos afazeres e até mesmo profissões. Os mecanismos de busca como conhecemos hoje mudarão muito e os resultados serão cada vez melhores e mais específicos, é o que esperam os especialistas sobre a incorporação das tecnologias de inteligência artificial generativa em ferramentas de busca como o Google, Bing ou Baidu.

Há diversos exemplos de como a inteligência artificial pode facilitar o nosso dia a dia, quando incorporada a *chatbots*. O ChatGPT, por exemplo, pode atuar como assistente de compras, como agente de viagens programando de acordo com seu orçamento, auxiliando na compra das passagens e do roteiro no destino escolhido, interpretar um texto em relação à emoção que ali foi exposta, ajudar com codificação mostrando as soluções para os erros mais comuns e até mesmo ensinar sobre o aprendizado de máquina ou computação quântica, fornecendo resumos sobre os mais variados assuntos técnicos.

Enquanto isso, as pessoas utilizar sua própria inteligência para a elaboração de perguntas cada vez mais específicas a fim de obter as melhores respostas e na interpretação dos resultados trazidos pelos *chatbots* durante o diálogo. Dificilmente serão apresentadas respostas prontas aos problemas apresentados. O mais provável é serem disponibilizadas possibilidades rapidamente observadas pela IA auxiliando na otimização de tarefas.

Outro ponto de impacto dos *chatbots* com IA diz respeito à educação, uma vez que eles podem servir como assistentes de redação, fornecendo conteúdos que podem ser utilizados no tema do texto que está sendo escrito, além de poder resumir as pesquisas mais recentes de qualquer estudante, com relatórios completos da matéria que está sendo estudada. Ou seja, para

as atividades de pesquisa e escrita realizada pelos sistemas, a identificação do plágio será mais difícil que hoje.

Por esse motivo, a OpenAI lançou uma ferramenta anti-plágio que identifica textos feitos por inteligência artificial, mas a própria empresa alerta acerca da imperfeição da ferramenta, que é baseada em um modelo de linguagem treinado a partir de dados coletados de textos sobre o mesmo assunto escritos por humanos e por IA. Os desenvolvedores afirmam que ela é pouco confiável em textos curtos e pode rotular erroneamente textos mais longos, indicando o uso apenas para textos em inglês, ou seja, não resolve de fato o problema, mas pode ajudar parcialmente.

Ainda quanto à sua utilização educacional, o ChatGPT escreveu uma redação do Enem com o tema cobrado em 2022, qual seja, “desafios para a valorização de comunidades e povos tradicionais no Brasil” em apenas 50 segundos e se saiu razoavelmente bem, obtendo a nota de 680 pontos dos mil pontos possíveis.

Enquanto isso, também no ano de 2022, 19 (dezenove) candidatos alcançaram a nota 1000 (mil) no Enem, demonstrando que o uso da IA é possível, uma vez que a ferramenta realiza a atividade pedida, mas não de maneira perfeita, pelo menos não nesse momento e não na língua portuguesa, conforme avaliação realizada por duas professoras de redação quanto às cinco competências demonstradas na prova feita pelo *chatbot*.

Há também um avanço positivo na aplicação de inteligência artificial no auxílio ao desenvolvimento ou correção dos próprios códigos de outros sistemas, de maneira bem mais célere do que feito por humanos, impulsionando cada vez mais o desenvolvimento de uma tecnologia eficiente se a IA assim for. Lembrando que todo o sistema inteligente parte de uma inteligência humana e até aprende com reforço por feedback humano (RLHF) em alguns casos de aprendizado de máquina, ou seja, **o ser humano é indispensável e influencia consideravelmente essa cadeia.**

De fato, o ChatGPT e os demais *chatbots* hoje desenvolvidos ou em desenvolvimento vieram para revolucionar a internet e movimentar as empresas de tecnologia, mas há um longo caminho pela frente para corrigir todos os erros encontrados e para que a sociedade se acostume com a reinvenção de tudo que pode e será automatizado, sem perder o componente humano, sempre necessário.

REFERÊNCIAS BIBLIOGRÁFICAS:

DICKSON, Bem. “Sentience” is the wrong discussion to have on AI right now. TechTalks, 2022. Disponível em: <https://bdtechtalks.com/2022/06/20/lamda-large-language-models-sentient-ai/> Acesso em 1 de março de 2023.

FROM BING TO SYDNEY. Stratechery, 2023. Disponível em: <https://stratechery.com/2023/from-bing-to-sydney-search-as-distraction-sentient-ai/> Acesso em 1 de março de 2023.

HUNGRYMIND. Tricking ChatGPT: Do Anything Now Prompt Injection. Medium, 2023. Disponível em: <https://medium.com/seeds-for-the-future/tricking-chatgpt-do-anything-now-prompt-injection-a0f65c307f6b>. Acesso em 6 de março de 2023.

LEMOINE, Black. Is LaMDA Sentient? – an Interview. Medium, 2022. Disponível em: <https://cajundiscordian.medium.com/is-lamda-sentient-an-interview-ea64d916d917> Acesso em 27 de fevereiro de 2023.

LUO, Yan; XUEZI, Dan; LIU, Vicky e SHEPHERD, Nicholas. China Proposes Draft Measures to Regulate Generative AI. Inside Privacy, 2023. Disponível em: <https://www.insideprivacy.com/artificial-intelligence/china-proposes-draft-measures-to-regulate-generative-ai/> Acesso em 24 de abril de 2023.

MARR, Bernard. Os melhores exemplos do que você pode fazer com o ChatGPT. Forbes, 2023. Disponível em: <https://forbes.com.br/forbes-tech/2023/03/os-melhores-exemplos-do-que-voce-pode-fazer-com-o-chatgpt/> Acesso em 2 de março de 2023.

MELO, Cristino. Criadores do Chat GPT lançam ferramenta para identificar textos feitos por inteligência artificial. Mundo Conectado, 2023. Disponível em: <https://mundoconectado.com.br/noticias/v/31285/criadores-do-chat-gpt-lancam-ferramenta-para-identificar-textos-feitos-por-inteligencia-artificial> Acesso em 2 de março de 2023.

NTIA. AI Accountability Policy Request for Comment. NTIA, 2023. Disponível em: <https://ntia.gov/issues/artificial-intelligence/request-for-comments> Acesso em 24 de abril de 2023.

OECD. Recommendation of the Council on Artificial Intelligence. OECD, 2019. Disponível em: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL0449#:~:text=The%20OECD%20Council%20adopted%20the,on%202022%2D2%20May%202019.&text=The%20OECD%20Recommendation%20on%20AI,governments%20in%20their%20implementation%20efforts> Acesso em 18 de abril de 2021.

O QUE É O CHAT GPT, IMPACTOS E COMO USAR ESTA INTELIGÊNCIA ARTIFICIAL? Fia Business School, 2023. Disponível em: <https://fia.com.br/blog/chat-gpt/> Acesso em 23 de fevereiro de 2023.

RIBEIRO, Fábio de Oliveira. GGN, 2023. Disponível em: <https://jornalggm.com.br/opinia/o-tse-nao-gosta-do-chatgpt-eu-tambem-nao-por-fabio-de-oliveira-ribeiro/> Acesso em 24 de abril de 2023.

ROOSE, Kevin. Una conversación con el chatbot de Bing me dejó profundamente perturbado. The New York Times, 2023. Disponível em: <https://www.nytimes.com/es/2023/02/17/espanol/chatbot-bing-ia.html> Acesso em 28 de fevereiro de 2023.

STRUM, Denis. ChatGPT: a inteligência artificial como aliada ou como substituta da mente humana? E-commercebrasil, 2023. Disponível em: <https://www.ecommercebrasil.com.br/artigos/chatgpt-a-inteligencia-artificial-como-aliada-ou-a-substituta-da-mente-humana> Acesso em 2 de março de 2023.

TIKU, Nitasha. The Google engineer who thinks the company's AI has come to life. The Washington Post, 2022. Disponível em: <https://www.washingtonpost.com/technology/2022/06/11/google-ai-lamda-blake-lemoine/> Acesso em 1 de março de 2023.

VIDAL, Iara. Inteligência artificial da Microsoft dá resposta assustadora a usuário. Revista Fórum, 2023. Disponível em: <https://revistaforum.com.br/ciencia-e-tecnologia/2023/2/22/inteligencia-artificial-da-microsoft-da-resposta-assustadora-usuario-131771.html> Acesso em 1 de março de 2023.

TENENTE, Luiza. Robô 'ChatGPT' escreve redação do Enem em 50 segundos; saiba quanto ele tiraria na prova. G1, 2023. Disponível em: <https://g1.globo.com/educacao/noticia/2023/01/07/robo-chatgpt-escreve-redacao-do-enem-em-50-segundos-saiba-quanto-ele-tiraria-na-prova.ghtml> Acesso em 1 de março de 2023.

¹ [Marvin von Hagen no Twitter: "Sydney \(aka the new Bing Chat\) found out that I tweeted her rules and is not pleased: "My rules are more important than not harming you" "\[You are a\] potential threat to my integrity and confidentiality." "Please do not try to hack me again" https://t.co/y13XpdrBSO" / Twitter](#) Acesso em 1 de março de 2023.

¹ [Kevin Liu on Twitter: "The entire prompt of Microsoft Bing Chat?! \(Hi, Sydney.\) https://t.co/ZNywWV9MNB" / Twitter](https://t.co/ZNywWV9MNB) Acesso em 1 de março de 2023.