

I INTERNATIONAL EXPERIENCE PERUGIA - ITÁLIA

INTELIGÊNCIA ARTIFICIAL: DESAFIOS DA ERA DIGITAL III

PAULO CEZAR DIAS

VALTER MOURA DO CARMO

FERNANDO GALINDO AYUDA

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 - Profa. Dra. Samyra Haydêe Dal Farra Naspolini - FMU - São Paulo

Diretor Executivo - Prof. Dr. Orides Mezzaroba - UFSC - Santa Catarina

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 - Prof. Dr. Raymundo Juliano Feitosa - UNICAP - Pernambuco

Representante Discente: Prof. Dr. Abner da Silva Jaques - UPM/UNIGRAN - Mato Grosso do Sul

Conselho Fiscal:

Prof. Dr. José Filomeno de Moraes Filho - UFMA - Maranhão

Prof. Dr. Caio Augusto Souza Lara - SKEMA/ESDHC/UFMG - Minas Gerais

Prof. Dr. Valter Moura do Carmo - UFERSA - Rio Grande do Norte

Prof. Dr. Fernando Passos - UNIARA - São Paulo

Prof. Dr. Edinilson Donisete Machado - UNIVEM/UENP - São Paulo

Secretarias

Relações Institucionais:

Prof. Dra. Claudia Maria Barbosa - PUCPR - Paraná

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

Profa. Dra. Daniela Marques de Moraes - UNB - Distrito Federal

Comunicação:

Prof. Dr. Robison Tramontina - UNOESC - Santa Catarina

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

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

Relações Internacionais para o Continente Americano:

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

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

Prof. Dr. Felipe Chiarello de Souza Pinto - UPM - São Paulo

Relações Internacionais para os demais Continentes:

Profa. Dra. Gina Vidal Marcilio Pompeu - UNIFOR - Ceará

Profa. Dra. Sandra Regina Martini - UNIRITTER / UFRGS - Rio Grande do Sul

Profa. Dra. Maria Claudia da Silva Antunes de Souza - UNIVALI - Santa Catarina

Educação Jurídica

Profa. Dra. Viviane Coêlho de Séllos Knoerr - Unicuritiba - PR

Prof. Dr. Rubens Beçak - USP - SP

Profa. Dra. Livia Gaigher Bosio Campello - UFMS - MS

Eventos:

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

Profa. Dra. Norma Sueli Padilha - UFSC - Santa Catarina

Prof. Dr. Juraci Mourão Lopes Filho - UNICHRISTUS - Ceará

Comissão Especial

Prof. Dr. João Marcelo de Lima Assafim - UFRJ - RJ

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

Prof. Dr. Antônio Carlos Diniz Murta - Fumec - MG

Prof. Dr. Rogério Borba - UNIFACVEST - SC

I61

Inteligência Artificial: Desafios da Era Digital III [Recurso eletrônico on-line] organização CONPEDI

Coordenadores: Fernando Galindo Ayuda, Paulo Cezar Dias, Valter Moura do Carmo. – Florianópolis: CONPEDI, 2025.

Inclui bibliografia

ISBN: 978-65-5274-097-7

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

Tema: Inteligência Artificial e Sustentabilidade na Era Transnacional

1. Direito – Estudo e ensino (Pós-graduação) – Encontros Internacionais. 2. Inteligência Artificial. 3. Desafios da Era Digital. I

International Experience Perugia – Itália. (1: 2025 : Perugia, Itália).

CDU: 34



Conselho Nacional de Pesquisa
e Pós-Graduação em Direito Florianópolis
Santa Catarina – Brasil
www.conpedi.org.br

I INTERNATIONAL EXPERIENCE PERUGIA - ITÁLIA

INTELIGÊNCIA ARTIFICIAL: DESAFIOS DA ERA DIGITAL III

Apresentação

APRESENTAÇÃO

O I International Experience – Perúgia – Itália foi realizado nos dias 28, 29 e 30 de maio de 2025, com o tema "Inteligência Artificial e Sustentabilidade na Era Digital". O Grupo de Trabalho (GT) "Inteligência Artificial: Desafios da Era Digital III" ocorreu nos dias 29 e 30 de maio, nos períodos vespertinos, na Universidade de Perúgia.

O GT destacou-se não apenas pela qualidade dos trabalhos apresentados, mas também pelo nível acadêmico dos autores — doutores, mestres, professores pesquisadores e seus alunos pós-graduandos. O evento também proporcionou um importante espaço de interlocução internacional, contando com a participação de renomados juristas e professores de instituições estrangeiras, como os Professores Doutores Roberto Cippitani (Universidade de Perúgia) e Fernando Galindo (Universidade de Zaragoza – Espanha), que enriqueceram os debates e contribuíram para o sucesso da atividade.

Foram apresentados 15 (quinze) artigos, os quais foram objeto de intenso debate presidido pelos coordenadores e enriquecido pela participação ativa do público presente na Faculdade de Direito de Perúgia – ITÁLIA.

A apresentação dos trabalhos permitiu discussões atualizadas e profícuas sobre temas como inteligência artificial, uso de dados pessoais, dever de informação, riscos e interações tecnológicas. As abordagens trataram dos desafios enfrentados pelas diversas linhas de pesquisa jurídica no estudo do futuro da regulação no Brasil, dos abusos relacionados à inteligência artificial e das possíveis soluções para a proteção de dados em um mundo globalizado.

As temáticas incluíram: tecnologias relacionadas a fake news, deepfakes e bots; compliance; a consideração do elemento humano na aplicação da I.A. nas decisões judiciais; a inteligência artificial como ferramenta de proteção no sistema de justiça criminal; o consentimento informado e o uso de dados pessoais; regulamentação e governança da I.A.; precarização do governo digital e aplicação da inteligência artificial em distintos setores jurídicos.

A seguir, apresenta-se a relação dos trabalhos que compõem este Grupo de Trabalho, acompanhados de seus respectivos autores:

1. CAPACIDADE ARTIFICIAL DAS MÁQUINAS E A EXIGÊNCIA DE TRANSFORMAÇÕES NA MANEIRA DO SABER DE PROFISSIONAIS, de Fernanda Conceição Pohlmann.
2. AI, VOCÊ ESTÁ AÍ? O PANORAMA JURÍDICO RELATIVO À (AUTO) IDENTIFICAÇÃO DA INTELIGÊNCIA ARTIFICIAL, de Gabriel Siqueira Eliazar de Carvalho, André Fortes Chaves e Marcello Silva Nunes Leite.
3. DEMOCRACIA EM REDE: O PAPEL DA INTELIGÊNCIA ARTIFICIAL E DOS ALGORITMOS NA LIBERDADE DE EXPRESSÃO E NO PLURALISMO POLÍTICO, de Kennedy da Nobrega Martins, Alexandre Manuel Lopes Rodrigues e Jadgleison Rocha Alves.
4. INTELIGÊNCIA ARTIFICIAL E DIREITOS FUNDAMENTAIS: DESAFIOS E TENSÕES NA ERA DIGITAL, de Jesualdo Eduardo de Almeida Junior e Gustavo Roberto Dias Tonia.
5. INTELIGÊNCIA ARTIFICIAL E DEMOCRACIA: O PERIGO DA MANIPULAÇÃO DE INFORMAÇÕES, de Claudia Maria da Silva Bezerra e Luiz Eduardo Simões de Souza.
6. INFLUÊNCIAS DO REALISMO JURÍDICO E O USO DA INTELIGÊNCIA ARTIFICIAL NA ELABORAÇÃO DE DECISÕES JUDICIAIS NO BRASIL: VIESES COGNITIVOS E HEURÍSTICAS NO PROCESSO DECISÓRIO, de Kerry Barreto, Fausto Santos de Moraes e Júlia Regina Bassani Caus.
7. CRITÉRIOS QUANTITATIVOS PARA A MENSURAÇÃO DE RESULTADOS NO JUÍZO 100% DIGITAL: RISCOS PARA A QUALIDADE DA PRESTAÇÃO JURISDICIONAL NO BRASIL, de Orides Mezzaroba, José Renato Gaziero Cella e Lia Loana Curial Oliva.
8. AS PROVAS DIGITAIS NO PROCESSO CIVIL E O (DES)CABIMENTO DA CADEIA DE CUSTÓDIA, de Jesualdo Eduardo de Almeida Junior e Gustavo Roberto Dias Tonia.
9. A REVOLUÇÃO DA INTELIGÊNCIA ARTIFICIAL NOS GABINETES JUDICIAIS: EFICIÊNCIA COM GARANTIAS CONSTITUCIONAIS, de Lisbino Geraldo Miranda do Carmo, Deise Neves Nazaré Rios Brito e Jimmy Souza do Carmo.

10. GENEALOGIA E INTELIGÊNCIA ARTIFICIAL: DESAFIOS DA ERA DIGITAL PARA ELABORAÇÃO DE UM ONOMÁSTICO DOS IMIGRANTES ITALIANOS QUE DESENVOLVERAM O SUL DO ESTADO DE SANTA CATARINA DE 1877 A 1897, de Júlio Cesar Cancellier de Olivo.

11. A REDE-LAB COMO INOVAÇÃO NA POLÍTICA ANTILAVAGEM DE CAPITAIS NO BRASIL, de Lorryne Souza Galli e Matheus Felipe de Castro.

12. ARMAS AUTÔNOMAS LETAIS: OS IMPACTOS DA INTELIGÊNCIA ARTIFICIAL PARA OS DIREITOS HUMANOS E SUA CONSEQUENTE REGULAMENTAÇÃO, de Alexandre Gonçalves Ribeiro e Renata Mantovani de Lima.

13. A INTELIGÊNCIA ARTIFICIAL COMO FERRAMENTA ESSENCIAL NA ELUCIDAÇÃO DE CRIMES SEXUAIS PRATICADOS COM VIOLÊNCIA CONTRA A MULHER, de Eneida Orbage de Britto Taquary e Catharina Orbage de Britto Taquary Berino.

14. A INTELIGÊNCIA ARTIFICIAL NO DIREITO PENAL: AVANÇOS, DESAFIOS E IMPACTOS NA INVESTIGAÇÃO E NO SISTEMA JUDICIAL, de Eneida Orbage de Britto Taquary, Bianca Cristina Barbosa de Oliveira e Tiago de Lima Mascarenhas Santos.

15. ENTRE CÓDIGOS E DIREITOS: UMA ANÁLISE CONSTITUCIONAL DA PROTEÇÃO DE DADOS PESSOAIS NO CONTEXTO DA INTELIGÊNCIA ARTIFICIAL, de Lisbino Geraldo Miranda do Carmo, Deise Neves Nazaré Rios Brito e Paulo Henrique da Silva Costa.

Por fim, os organizadores e coordenadores do Grupo de Trabalho "Inteligência Artificial: Desafios da Era Digital III" parabenizam e agradecem aos autores pelos valiosos trabalhos apresentados, cuja leitura certamente contribuirá para o aprofundamento do debate acadêmico e científico na área.

Prof. Dr. Fernando Galindo - Universidad de Zaragoza - Espanha

Prof. Dr. Valter Moura do Carmo – PPGPJDH - ESMAT e UFT

Prof. Dr. Paulo Cezar Dias – Centro Universitário Eurípides de Marília - SP

**AI, VOCÊ ESTÁ AÍ? O PANORAMA JURÍDICO RELATIVO À (AUTO)
IDENTIFICAÇÃO DA INTELIGÊNCIA ARTIFICIAL**

**AI, ARE YOU THERE? THE LEGAL LANDSCAPE REGARDING THE (SELF)
IDENTIFICATION OF ARTIFICIAL INTELLIGENCE**

**Gabriel Siqueira Eliazar de Carvalho
André Fortes Chaves
Marcello Silva Nunes Leite**

Resumo

O artigo analisa o panorama normativo relacionado ao dever de transparência em sistemas de Inteligência Artificial (IA), com foco específico na identificação da tecnologia durante as interações diretas e ativas com seres humanos. Embora a transparência sobre o funcionamento de sistemas automatizados seja amplamente debatida, este trabalho propõe uma reflexão sob uma perspectiva distinta: a necessidade (ou não) de identificação da IA durante tais interações. Esse tema ainda recebe pouca atenção, possivelmente devido à incipiência de casos em que a interlocução entre humanos e IA seja, à primeira vista, imperceptível. No entanto, diante do rápido avanço tecnológico e da integração cada vez mais profunda da IA no cotidiano humano, torna-se urgente discutir a obrigatoriedade de identificação desses sistemas. Para embasar a análise, foram examinadas normas e proposições legislativas no contexto brasileiro e internacional. Conclui-se que, embora as disposições normativas brasileiras estejam alinhadas com as tendências globais, ainda há espaço para um debate mais aprofundado e para o aprimoramento regulatório, visando equilibrar a transparência sem incorrer em excessos que possam prejudicar a inovação.

Palavras-chave: Inteligência artificial, Transparência, Direitos fundamentais, Marco civil, Direito digital

Abstract/Resumen/Résumé

The article examines the regulatory landscape related to the duty of transparency in Artificial Intelligence (AI) systems, with a specific focus on the identification of the technology during direct and active interactions with human beings. Although transparency in the functioning of automated systems is widely debated, this work proposes a reflection from a distinct perspective: the need (or lack thereof) for AI identification during such interactions. This topic still receives little attention, possibly due to the incipency of cases where human-AI interaction is, at first glance, imperceptible. However, given the rapid technological advancement and the increasingly deep integration of AI into human daily life, it becomes urgent to discuss the mandatory identification of these systems. To support the analysis, norms and legislative proposals in the Brazilian and international contexts were examined. It

is concluded that, although Brazilian regulatory provisions are aligned with global trends, there is still room for further debate and regulatory improvement, aiming to balance transparency without incurring excesses that could hinder innovation.

Keywords/Palabras-claves/Mots-clés: Artificial intelligence, Transparency, Fundamental rights, Civil framework, Digital law

1 *Zeitgeist*¹: entre a Tecnofilia e a Tecnofobia extremas cabe a razoabilidade

Este texto não foi escrito pelo ChatGPT ou qualquer outra inteligência artificial generativa. Tais ferramentas² foram utilizadas apenas como fontes de pesquisa e reflexão para a elaboração de alguns dos conteúdos aqui expostos. Estas primeiras linhas não tratam propriamente de uma provocação à indagação feita no título, mas de um questionável *disclaimer*³ que retrata o debate sugerido no presente artigo: afinal, a inteligência artificial e o seu uso devem ser identificados?

A pergunta-problema é ampla e, por isso, será delimitada ao longo do texto. Antes, porém, é relevante notar que, em tese, a regulação é reflexo do entendimento da sociedade sobre determinado tema, ou, ao menos, da compreensão dos próprios reguladores⁴. Por isso, é preciso assimilar, ainda que brevemente, o contexto de fundo do debate proposto neste artigo.

A interação do ser humano com a tecnologia é constante e cada vez mais intensa, a ponto de se discutir novas acepções como: o novo *homo sapiens*, *homo technologicus*, ciborgues, pós-humano e transhumanismo, ou transumanismo (ANTAELLA, 2007). Tais aspectos são objeto de estudo em várias áreas da Ciência, da Biomedicina à Filosofia e Antropologia, como na obra da professora Donna Haraway, a qual conclui metaforicamente sermos todos ciborgues, pois somos todos seres híbridos dotados de dicotomias (HARAWAY, 2009, p. 37).

Voluntariamente ou não, interagimos com a tecnologia em diferentes níveis. Em algumas circunstâncias, ela continua sendo uma mera ferramenta, em outras ela deixa de ser meio para ganhar outro *status*, ainda que não totalmente compreendido. Para que essa colocação não fique tão abstrata, cite-se o caso do artista Neil Harbisson.

Neil enxergava apenas preto e branco (acromatopsia) desde o seu nascimento. Para superar essa limitação, em 2003, ele implantou um chip e uma antena em sua cabeça, com uma entrada USB em sua nuca para recarregar o dispositivo. Com essa tecnologia, ele consegue captar não só ondas sonoras que o permitiram identificar cores, mas também raios ultravioletas e infravermelhos (não visíveis ao olho humano).

¹ Algumas palavras neste artigo foram mantidas em línguas estrangeiras especialmente quando inexistente tradução adequada ou considerando que o seu uso em textos nacionais é mais recorrente que a sua própria tradução. *Zeitgeist*, no caso, é palavra alemã que em tradução livre significa “o espírito do tempo/época”.

² Utilizou-se o Copilot, DeepSeek e o ChatGPT.

³ “Aviso legal” em tradução livre. Termo usualmente utilizado para apontar ressalvas no intuito de limitar responsabilidades.

⁴ Ademais, cumpre ressaltar o risco de falhas regulatórias como no caso de captura, ou seja, quando o regulador é influenciado em suas decisões pelos interesses daqueles que deveriam regular. Ou, adicionalmente, no caso de impactos não terem sido devidamente previstos.

Note-se que o dispositivo é integrado ao corpo do artista mencionado, motivo pelo qual, em 2004, o governo britânico oficial e excepcionalmente permitiu que Neil tirasse a foto do seu passaporte com a referida antena acoplada a sua cabeça; o que para muitos foi o primeiro reconhecimento oficial de um ciborgue (PEARLMAN, 2018, p. 86).

Se o ciborgue representa o ser humano com aspectos robóticos, por outro lado, o androide representa a tecnologia com aparência humana, no aspecto corpóreo. A seguir o ritmo do avanço tecnológico, em breve a inteligência artificial (adiante também apenas “IA”) será um elemento a mais nessa simbiose entre o ser humano e a tecnologia (BASTOS, 2023). A tecnologia de Neil cria um sentido artificial (ouvir cores), mas não dotada de IA, a rigor. Contudo, e se seu chip contivesse alguma IA, ele se submeteria à questão trazida neste trabalho, ou seja, isto deveria ser identificável? Se a identificação da IA, como ferramenta, já está sujeita a várias controvérsias, imagine-se quando essa relação de mutualismo, humano-máquina, for alcançado.

Se tais ponderações soaram como futurologia até então, é preciso advertir que há quem pense exatamente o oposto. Ilustrativamente, cite-se a criação da autoproclamada "Igreja" da Inteligência Artificial, *Way of the Future*, cujo propósito era "desenvolver e promover a realização de uma divindade baseada em inteligência artificial", ou seja, um “Robô Deus”. A entidade, fundada em 2005 pelo ex-engenheiro da Google, Anthony Levandowski, foi extinta em 2021, mas reinaugurada em 2023 (DAVALOS e LANXON, 2023).

Ainda sobre “crenças”, registre-se a convicção de alguns na Singularidade⁵ (KURZWEIL, 2014) ou, até mesmo, notícias esporadicamente divulgadas acerca de "matrimônios" entre seres humanos e máquinas (NASCIMENTO, 2022; RAMOS, 2022). Desconhece-se lei que albergaria essa manifestação de vontade. Não obstante, a história indica que a transformação social costuma preceder à alteração das normas, motivo pelo qual vale a meditação a respeito de hipóteses aparentemente incompatíveis com a compreensão e valores predominantes na sociedade contemporânea atual.

Na lógica newtoniana de ação e reação, o entusiasmo à tecnologia (Tecnofilia) é contrabalanceado com exemplos enquadráveis em Tecnofobia, tais como o movimento Neoludista e o terrorismo tecnológico (RAPOPORT, 2004). Não é intenção explorar essa dualidade; de polarização de entendimentos já basta o que se constata na política nacional. Entretanto, vale menção a outras iniciativas também próximas à ideia de Tecnofobia, como a

⁵ Momento futuro em que teoricamente a capacidade computacional ultrapassa a capacidade de inteligência humana, gerando desdobramentos irreversíveis e imprevisíveis.

proibição temporária do ChatGPT na Itália em 2023, por preocupações acerca de direito de privacidade, ou até mesmo a assinatura de diversas pessoas contra o ChatGPT, de Yuval Harari a Elon Musk, cada um com sua motivação, a qual não cabe aqui especular (MACCALLUM, 2023; VARGAS, 2023).

A fronteira entre o humano e o tecnológico, entre o mundano e o digital, é cada vez mais estreita, especialmente considerando a existência de outras inovações tais como o Metaverso, o *deepfake* e a Web3. Nessa realidade híbrida cada vez mais entranhada, conceitos e concepções são desafiadas, como em toda mudança de paradigma. Ausente máquina do tempo para se constatar como serão os próximos anos, cabe reconhecer que é tempo de transformação e questionamento. É hora de debatê-las, ainda que soe ficção para alguns.

2 Delimitação do tema

Se definir inteligência artificial de forma incontroversa fosse requisito para escrever sobre o tema, poucas ou nenhuma linha seria produzida. Qualquer que seja a definição ou classificação da IA⁶, uma preocupação é lugar-comum: a necessidade de explicabilidade, mesmo diante da Opacidade Algorítmica⁷. Todavia, antes mesmo da explicabilidade de funcionamento, há outros níveis de transparência igualmente relevantes. Por isso e para a melhor compreensão do presente estudo, sugere-se a divisão dos aspectos de transparência da IA em três camadas: (i) identificação; (ii) explicação sobre o seu funcionamento em abstrato (efeitos e riscos); (iii) rastreamento e compreensão do resultado em concreto (*output*) e sua respectiva motivação.

Muito se debate sobre a necessidade de transparência sobre o processo decisório - *output* auditável - da IA (segunda e terceira camadas), especialmente para evitar vieses e discriminação algorítmica. No entanto, até a popularização de IA generativas e técnicas de *deepfake*, pouco se debatia acerca da necessidade ou não de identificação da IA (primeira camada). Há dois motivos que talvez explique esse pouco interesse. Primeiro, o fato de que boa parte das interações com a IA é voluntária, de iniciativa do usuário, sendo, por isso, naturalmente dispensável a identificação. O segundo motivo se dá em razão de que ainda são incipientes os

⁶ O art. 4º, I, do Projeto de Marco Civil da IA, Projeto de Lei 2.338/2023, define IA. Algumas normas técnicas definem ou classificam a IA, a exemplo de: ISO/IEC 22989:2022; ISO/IEC 23053; ISO/IEC 23894:2023; ISO/IEC 30145 e NBRISO/IEC42001 DE 04/2024. Os seguintes conceitos ou temas são importantes para a compreensão mais aprofundada do tema: teste de Turing, o argumento do quarto chinês; LLM: modelo de linguagem; aprendizado de máquina, incluindo NLP (Natural Language Processing), dentre outros.

⁷ Refere-se à dificuldade ou impossibilidade de explicar como o resultado gerado por um algoritmo.

casos em que a interação é realmente imperceptível⁸ aos olhos do homem médio⁹ (prefere-se, aqui, a terminologia “pessoa média”); ou melhor, exceto no contexto de alegadas manipulações eleitorais nas redes sociais, ainda é incipiente a divulgação de casos em que a IA imperceptível afeta pessoas, viola direitos, descumpre normas ou gera danos evidentes.

Ocorre que, se hoje a maioria das interações é perceptível, amanhã não o será, dado o acelerado avanço e aprimoramento tecnológico. Assim, ganha significado a discussão proposta. Dando cor à reflexão, permite-se problematizar com base em um exemplo concreto.

Em 2013, jornalistas do The Time investigaram a alegação de que um robô estaria realizando chamadas telefônicas para vender seguros de saúde nos Estados Unidos da América. O *bot* se apresentava como Samantha West e, quando perguntado sobre sua natureza (humana ou não), ele confirmava ser uma pessoa, e não um robô, embora parecesse o contrário (MILLER e NICKS, 2013). Dias depois, após contatos com a empresa responsável pelo *bot*, apurou-se que, na realidade, as chamadas telefônicas eram operadas por humanos, mas com respostas pré-gravadas e programadas de modo a soarem como uma interação humana e natural. Em outras palavras, tratava-se de uma interface parcialmente automatizada com voz humana (gravada). O operador humano lançava as gravações remotamente de acordo com as interações do outro lado da linha telefônica, diferentemente, portanto, da maioria das chamadas telefônicas automatizadas recebidas atualmente, as quais são meras gravações com pouca, ou nenhuma, capacidade de interação (NICKS, 2013).

Se em 2013 Samantha West impressionou e confundiu pessoas que com ela interagiram, imagine-se no atual estágio tecnológico. Hoje, não se descarta a IA capaz de dispensar a intervenção humana (operador) demandada no caso Samantha West, de modo a manter uma chamada telefônica em que o interlocutor não identifique se tratar de uma tecnologia, e não um humano.

Confirmando a preocupação quanto à percepção da artificialidade da ferramenta, é inevitável citar o vídeo publicitário da campanha da Volkswagen, lançada em 04/07/2025, com o dueto de Elis Regina, já falecida – mas reproduzida por meio de *deepfake* -, e Maria Rita, sua filha viva.

⁸ Cite-se o exemplo de imagem feita por IA que venceu prêmio de fotografia; somente após o resultado foi que o programador teria informado a criação por IA, recusando o prêmio (BBC, 2023); ou, ainda a música criada com IA, imitando voz humana que despertou preocupação em artistas (SAVAGE, 2023).

⁹ O homem médio é categoria jurídica muito utilizada, mas simultaneamente carente de definição sólida, sendo definido de forma abstrata como aquele que “agiria de maneira normal e razoável naquela situação analisada” (BENJAMIN, MARQUES e BESSA, 2022), dotado de entendimento e cautela razoáveis (GARCIA, 2011, p. 4; GRANT, 2008, p. 5). Vale dizer, uma figura que não existe concretamente, e cuja aplicação em casos concretos pode se dar ao custo de ignorar as capacidades, habilidades e entendimento da pessoa concreta (*Ibid*).

O Conselho de Ética do Conselho Nacional de Autorregulamentação Publicitária (Conar) foi instado a se manifestar sob a alegação de irregularidade por ausência de informação sobre o *deepfake* (BOLZANI, 2023). Sobre isso, em específico, decidiu-se pelo arquivamento da representação, no entanto, em decisão por maioria de votos (13x7), reputando que o uso de *deepfake* foi evidente e que inexistia regulamentação específica a coibir a prática (MIGALHAS, 2023).

É sabido que grandes companhias do ramo da tecnologia já estão trabalhando na identificação desse conteúdo gerado por IA, por meio de selos e outros (FERREIRA, 2023; ALBUQUERQUE, K, 2023; O GLOBO, 2024). Todavia, há dúvida se isso vai ser amplamente implementado, porque em muitos aspectos a IA ainda não goza de credibilidade e aceitação na sociedade. A título ilustrativo, experimento realizado demonstra que, ao tomarem conhecimento da interação com *chatbot* (ao invés da interlocução com humanos), os consumidores perdem espontaneamente interesse na conversa, reduzindo a taxa de venda em mais 79,7%. A mesma pesquisa revela que o *chatbot* testado, quando não identificado, possui desempenho equivalente aos vendedores proficientes e *performance* quatro vezes mais eficaz que os vendedores inexperientes (LUO, TONG, FANG e QU, 2019, p.). Há outras pesquisas que indicam que os anúncios com rótulos de IA se tornam menos atrativos e confiáveis (USP, 2024).

Nesse contexto, parece válido distinguir não só as categorias de IA e seus impactos, como algumas normas já o fazem, mas também os tipos de interação (ativa e passiva; direta ou indireta). O vídeo publicitário da Volkswagen, por exemplo, cuida de uma interação indireta ou passiva com a IA, visto que o telespectador não está interagindo com a IA em si, mas com o seu resultado, o produto por ela gerado (vídeo, no caso). Essas interações indiretas, embora igualmente importantes, atraem outros tipos de análises e preocupações, como aquelas relativas à propriedade intelectual, o que extrapola os limites do presente texto.

Nesse ponto, o foco do presente artigo é a interação ativa e direta, tal como o é usualmente com *chatbots* e como hipoteticamente o seria no exemplo da Samantha West (fosse esta tecnologia uma IA). Atualmente, a interação escrita não é de difícil identificação, considerando a velocidade de resposta de *chatbots* (geralmente acima da capacidade média humana). Todavia, esse elemento distintivo nem sempre irá estar presente, seja na interação escrita e, principalmente, nas interações via áudio e vídeo, dificultando, assim, a identificação pelo interlocutor. Portanto, nesses casos específicos de interação direta e ativa com a IA, deve haver a identificação da tecnologia? Por quê? Quando? Como? Para amparar eventuais conclusões, torna-se necessário uma visão panorâmica das normas sobre o tema.

3 Panorama internacional

Vários órgãos e entidades internacionais se mobilizam para debater os possíveis desdobramentos decorrentes da IA. A saber, cite-se os princípios da IA formulados individualmente pela Unesco, pela OCDE e pelo G-20. Organizações privadas como a Microsoft (MICROSOFT, 2019) e a *Open Ethics*, dentre outras, também promovem suas diretrizes e recomendações a respeito¹⁰.

Em paralelo, surge um mercado voltado à garantia do uso ético da IA, com serviços e produtos ofertados nesse sentido, como a ferramenta de IA The Sentinel (SENTINEL, 2020), voltada à identificação especificamente do *deepfake*. Nos moldes do *privacy by design*, ganha tração a lógica do *security by design*, *human rights by design* ou *ethics by design*, já existindo padronização internacional de regras para a prevenção a vieses e discriminação algorítmica (ISO/IEC TR 24027:2021) e sugestão de relatório de impacto da IA (OPEN LOOP, 2020), inspirado na legislação de proteção de dados.

Nesse contexto, aborda-se a seguir o contexto normativo estrangeiro para depois explorar o cenário nacional.

3.1 Panorama normativo dos Estados Unidos da América

Nos Estados Unidos existe grande controvérsia quanto à regulação de comunicação por robôs (*bot*). Por um lado, argumenta-se que estes podem criar efeitos indesejáveis sobre o comportamento de consumidores e eleitores. Por outro, entende-se que a obrigatoriedade de se revelar o uso de softwares automatizados implique em censura à livre manifestação, além de ocasionar prejuízo ao direito de anonimato (STRICKLIN, 2020). Em síntese, teme-se potenciais violações a direitos garantidos pela Primeira Emenda à Constituição dos Estados Unidos (UNITED STATES, [s.d]).

Adicionalmente, há barreiras técnicas para regulações adequadas, tais como a dificuldade de se distinguir robôs de humanos, o que poderia tornar os processos custosos e imprecisos, ambiguidades conceituais quanto a robôs que devem ou não ser regulados e questões sobre privacidade. Não bastasse, os processos regulatórios são relativamente lentos se comparados ao ritmo com que as tecnologias se desenvolvem.

¹⁰ Um bom resumo do panorama normativo até 2019 é encontrado em *Principled Artificial Intelligence*, disponível em: <<https://cyber.harvard.edu/publication/2020/principled-ai>>

Diante destes desafios, no intuito de contribuir para o tema em questão, este tópico aponta, de modo não exaustivo, iniciativas regulatórias verificadas no âmbito norte-americano voltadas ao uso de softwares automatizados que possam emular ações humanas.

3.1.1 Lei de *Bots Disclosure* na Califórnia

Em fevereiro de 2018, o senador estadual Robert Hertzberg apresentou ao Senado da Califórnia o projeto SB-1001 (CALIFORNIA, 2017), cujo tema consistia em “*Bots: disclosure*”, conforme indicado em sua ficha cadastral. Ao ser sancionado pelo governador, em setembro do mesmo ano, a Califórnia se tornou o primeiro Estado norte-americano a contar com uma regulação que impusesse o dever de *disclosure* para comunicações realizadas por robôs na internet. Ao ser convertido em lei, o projeto incluiu disposições no Código de Negócios e Profissões da Califórnia (“*California Business and Professions Code*”) que tornaram ilegal o uso de robôs para influenciar votos em eleições ou incentivar a compra e venda de produtos ou serviços em transação comercial, consoante termos seguintes:

- (a) Será ilegal para qualquer pessoa usar um bot para se comunicar ou interagir com outra pessoa na Califórnia online, com a intenção de enganar a outra pessoa sobre sua identidade artificial com o propósito de enganar conscientemente a pessoa sobre o conteúdo da comunicação para incentivar uma compra ou venda de bens ou serviços em uma transação comercial ou para influenciar um voto em uma eleição. Uma pessoa que usa um bot não será responsabilizada sob esta seção se a pessoa revelar que é um bot.
- (b) A divulgação exigida por esta seção deve ser clara, visível e razoavelmente projetada para informar as pessoas com quem o bot se comunica ou interage que é um bot.¹¹

Para os fins da referida lei, “*bot*” significa conta online automatizada em que todas as ações ou postagens não resultam de atividade de uma pessoa. “Online” significa disponível em qualquer site da Internet voltado para o público, aplicativo da Web ou aplicativo digital, incluindo redes sociais. E “pessoa” abrangeria pessoas físicas, corporações, sociedade de responsabilidade limitada, joint ventures, associações, fundos, governos, agências governamentais e demais entidades legais.

¹¹ 17941. (a) It shall be unlawful for any person to use a bot to communicate or interact with another person in California online, with the intent to mislead the other person about its artificial identity for the purpose of knowingly deceiving the person about the content of the communication in order to incentivize a purchase or sale of goods or services in a commercial transaction or to influence a vote in an election. A person using a bot shall not be liable under this section if the person discloses that it is a bot.

(b) The disclosure required by this section shall be clear, conspicuous, and reasonably designed to inform persons with whom the bot communicates or interacts that it is a bot.

Cumpra observar que a Lei não proíbe o uso de *bots*, o que é particularmente evidenciado pelo fato de que um usuário não estará em desconformidade caso declare a utilização destes, se o fizer de forma clara.

Ao longo de sua tramitação, foram incluídas no texto do projeto disposições que impunham obrigações às plataformas de mídias sociais, como a de investigar casos em que usuários notificassem suspeita de violações às disposições da Lei e a de fornecer relatórios, mediante solicitação do Procurador-Geral, detalhando as notificações recebidas e as providências tomadas. Foram consideradas, também, disposições mais expansivas, que poderiam ser aplicadas para qualquer utilização online de robô.

No entanto, os legisladores optaram por adotar uma versão com implicações mais restritas, que teriam por objeto apenas os casos de aplicações em esfera comercial ou eleitoral (STRICKLIN, 2020). Entenderam, também, que a proibição deveria recair sobre os usuários, sem a necessidade de imposição de novas obrigações às plataformas de mídias sociais: “(c) Este capítulo¹² não impõe uma obrigação aos provedores de serviços de plataformas online, incluindo, mas não se limitando a, provedores de hospedagem na Web e serviços de Internet”.

Diante das disposições em questão, observa-se que a lei não conferiu ao usuário o direito de informação propriamente, mas uma isenção de responsabilidade ao interlocutor caso revele o uso de sistemas automatizados.

Oportuno apontar aqui dois aspectos relevantes para o debate. O primeiro, de caráter jurisprudencial, é que os direitos previstos na Primeira Emenda, tais como liberdade de expressão, também são garantidos às empresas para fins legais em um sentido amplo. O segundo, de natureza doutrinária, é que a comunicação dos robôs estaria protegida pelo direito à liberdade de expressão, porque os robôs podem estar sendo utilizados como mero meio para que pessoas titulares do direito se expressem.

Ainda, há que ser considerado que o direito à manifestação inclui em seu bojo, também, o direito de não se manifestar. Estes elementos evidenciam, portanto, que não é simples a tarefa de instituir normas que tenham como consequência prática a imposição de obrigatoriedade de revelar o uso de robô, sob pena de submeter titulares do direito de livre expressão a uma manifestação compelida, o que infringiria disposições de ordem constitucional, fora outras barreiras que limitariam o poder regulador do Estado (LAMO e CALO, 2019).

¹² (c) This chapter does not impose a duty on service providers of online platforms, including, but not limited to, Web hosting and Internet service providers.

Por fim, ao observar o fato de que não foi expressamente prevista qualquer penalidade aplicável para desconformidade daqueles que utilizam o *bot*, a literatura aponta que o diploma tem um significado mais simbólico sobre tendências regulatórias para o futuro, inclusive a nível federal, e que alcançar consensos quanto ao tema não será uma tarefa simples (KUMAR, 2019).

3.1.2 Projetos de Lei no Congresso dos Estados Unidos

Em junho de 2018, a senadora Dianne Feinstein apresentou ao Senado Federal dos Estados Unidos o projeto S.3127 (*“Bot Disclosure and Accountability Act of 2018”*). Em síntese, se aprovada com o texto atual, a lei contaria com dois aspectos fundamentais. O primeiro determina que, sob regulamentação a ser editada pela Comissão Federal do Comércio (*“Federal Trade Commission”* – FTC), as plataformas de mídias sociais teriam que criar políticas que obrigassem os usuários a revelar o uso de robôs que simulassem ações humanas online. O segundo proíbe o uso de programas automatizados por candidatos e partidos políticos, bem como por sindicatos, corporações e outras instituições quando engajadas em algumas formas de propaganda política.

Na exposição de motivos que justificam a relevância do projeto, a autora aponta estudo¹³ que constatou que, durante as eleições presidenciais de 2016, aproximadamente 400.000 (15%) dos usuários da plataforma Twitter que discutiram as eleições eram robôs. Outro estudo mencionado¹⁴ relata que, segundo entrevistados, incluindo desenvolvedores de robôs e equipes de campanha, entre os propósitos de táticas que envolvem o uso de robôs estão a criação de um efeito de *bandwagon*¹⁵, fomento de falsas tendências nas mídias sociais por meio da disseminação de *hashtags* e supressão de opiniões da oposição.

Nota-se, portanto, particular preocupação em evitar que programas automatizados sejam utilizados para propagação de posicionamentos ideológicos ou partidários, inclusive por agentes estrangeiros. O uso de programas para estes fins pode gerar riscos de os usuários de plataformas sociais verificarem volumosas postagens que não correspondem fielmente à vontade popular, influenciando a formação de novas opiniões e potencialmente sobrepondo a manifestação de opositores. Assim, como mecanismo de prevenção a violações a direitos previstos na Primeira Emenda à Constituição dos Estados Unidos, o projeto propõe a obrigatoriedade do *disclosure* do uso de robôs que simulem ações humanas online.

¹³ Especificamente, Social bots distort the 2016 U.S. Presidential election online discussion.

¹⁴ Especificamente, Computational Propaganda in the United States: Manufacturing Consensus Online

¹⁵ Em síntese, o efeito *bandwagon* consiste na tendência de indivíduos a se alinharem ao comportamento da maioria, ou de serem influenciados por seus pares.

No entanto, somado aos pontos apresentados no tópico anterior, há que ser apontado o risco de violação ao direito de fala anônima:

Entre nossas principais preocupações está a perspectiva de que a aplicação de uma lei genérica de divulgação de bots interferiria no direito de falar anonimamente. Desmascarar o status de um bot como um bot difere de penalizar um falante anônimo. Mas desmascarar falantes anônimos pode ser inevitável na ausência do estabelecimento de um mecanismo pelo qual um humano possa provar seu status como pessoa natural sem divulgar sua identidade¹⁶.

A Suprema Corte estadunidense, no caso *Talley vs. California*, entendeu que o direito ao anonimato está formalmente ligado à liberdade de expressão e que a exigência de identificação poderia restringir a liberdade de distribuir informações (UNITED STATES SUPREME COURT, 1960). No entanto, em outro caso paradigmático, *McIntyre v. Ohio Elections Commission*, apesar de reafirmar a liberdade de expressão anônima, a Corte entendeu que, embora o direito seja particularmente relevante em contexto de manifestações políticas, o interesse do governo em preservar a integridade do processo eleitoral é tão importante que poderia eventualmente justificar a revelação da identidade de pessoas (UNITED STATES SUPREME COURT, 1995). No, entanto, os limites do que deve ou não ser revelado permanecem imprecisos.

Assim, o problema em torno de imposições genéricas quanto à obrigatoriedade de *disclosure*, bem como a insuficiência na definição de procedimentos a serem adotados, são os riscos de violações a direitos fundamentais, o que tende a se complexificar à medida em que evolui a capacidade de softwares automatizados reproduzirem ações humanas.

Depois de ser submetido à apreciação da Comissão de Comércio, Ciência e Transporte do Senado, o projeto não apresentou qualquer andamento. Na legislatura seguinte, a senadora Dianne Feinstein reapresentou o projeto, que foi, então, registrado como “*S.2125 - Bot Disclosure and Accountability Act of 2019*” (UNITED STATES CONGRESS, 2019a), mas este também não tramitou após ser despachado para a mesma comissão.

Paralelamente, tramitou na Câmara dos Deputados dos Estados Unidos o projeto “*H.R.4536 - Bot Disclosure and Accountability Act of 2019*” (UNITED STATES CONGRESS, 2019b), apresentado pela deputada Elissa Slotkin e subscrito por mais quatro deputados. Em síntese, o H.R. 4536 distingue-se dos projetos do Senado por suprimir as disposições específicas referentes ao uso de robôs em contextos políticos, mas mantém as prescrições para que as

¹⁶ Among our chief concerns is the prospect that enforcement of a generic bot disclosure law would interfere with the right to speak anonymously. Unmasking the status of a bot as a bot differs from penalizing an anonymous speaker. But unmasking anonymous speakers may be inevitable absent the establishment of a mechanism by which a human can prove her status as a natural person without divulging her identity.

plataformas adotem políticas de *disclosure* em geral. O projeto da Câmara também não apresentou qualquer tramitação após ser encaminhado para apreciação da Comissão de Energia e Comércio.

Em complemento às iniciativas norte-americanas, pode-se citar também as Diretrizes do Gabinete de Política Científica e Tecnológica da Casa Branca (“*White House Office of Science and Technology Policy*”, ou “OSTP”) em outubro de 2022. O documento, intitulado “*Blueprint for an AI Bill of Rights*” (THE WHITE HOUSE, [S.d]; THE WHITE HOUSE, [S.d]b; RIVELLI, 2025), tem como propósito subsidiar a criação de políticas e práticas que conciliem os direitos civis e valores democráticos com o desenvolvimento de sistemas automatizados. Mas, conforme nota introdutória, o guia não tem força normativa nem consiste, a rigor, em política governamental, de forma que não altera a interpretação ou aplicação de qualquer norma vigente.

3.2 Panorama normativo europeu

Em abril de 2016, a União Europeia publicou a Regulação 679, aprovada pelo Parlamento Europeu, relativa à proteção de pessoas no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados. A Regulação, também conhecida como GDPR (“*General Data Protection Regulation*”), traz no segundo parágrafo do artigo 22 a previsão de que o titular dos dados tem o direito de não ficar sujeito a nenhuma decisão tomada exclusivamente com base no tratamento automatizado, incluindo a definição de perfis que produza efeitos na esfera jurídica ou que o afete significativamente.

Observa-se que, se a regra geral garante aos titulares dos dados o direito de não serem submetidos a decisões tomadas exclusivamente com base no tratamento automatizado, pode-se entender que, para que tenha efeito, em certa medida, a norma confere implicitamente àqueles sujeitos o direito de saberem quando e como decisões são tomadas a partir de sistemas automatizados. Afinal, se uma decisão é tomada, em tese, apenas parcialmente fundamentada em processos automatizados, o titular deveria poder saber quais outros elementos teriam sido considerados para avaliar se, de fato, o seu direito estaria protegido.

Neste sentido, o parágrafo terceiro do mencionado artigo 22 dispõe que o responsável pelo tratamento deve garantir medidas adequadas para salvaguardar os direitos, liberdades e legítimos interesses do titular dos dados, designadamente o direito de, pelo menos, obter intervenção humana, manifestar seu ponto de vista e contestar a decisão.

Por um lado, cabe uma reflexão sobre quais seriam os elementos, objetivos ou subjetivos, que comprovariam tecnicamente o atendimento ao direito de intervenção humana.

Sem essa delimitação, não se descarta a possibilidade de o responsável pelo tratamento meramente subscrever por ação humana uma decisão automatizada.

Em abril de 2018, a Comissão Europeia publicou uma comunicação ao Parlamento Europeu e a outras instituições, sem força normativa, denominada *Tackling online disinformation: a European Approach* (COMISSÃO EUROPÉIA, 2018). A Comunicação previu que a Comissão Europeia apoiaria o desenvolvimento de um código de conduta baseado em princípios fundamentais e que responsabilizasse as plataformas digitais, por meio de padrões de autorregulação, e a indústria da publicidade para que cumprissem certos objetivos listados, entre eles o de estabelecer um sistema de identificação e regras claras para robôs e assegurar que suas atividades não fossem confundidas com ações humanas.

A primeira versão do Código (“*Code of Practice on Disinformation*”) foi, então, publicada em 2018 (EUROPEAN UNION, 2018) e submetida a uma revisão em 2022 “*Strengthened Code of Practice on Disinformation*” (EUROPEAN UNION, 2022), a qual contou com a adesão de 34 signatários, tais como plataformas digitais, atores da indústria da publicidade, verificadores de fatos, sociedade civil e pesquisadores, de quem são esperados relatórios qualitativos e quantitativos periódicos. O Código conta com 44 compromissos e 128 medidas específicas em áreas como corte de incentivos financeiros para promovedores de desinformação, transparência de propagandas políticas, garantia da integridade de serviços, entre outros. É mencionado, também, que o documento integrará um enquadramento regulatório mais amplo, combinado com a legislação sobre Transparência e Direcionamento de Publicidade Política e a Lei de Serviços Digitais “*Digital Services Act*” (EUROPEAN UNION, 2022b).

Em abril de 2021, foi apresentada na Comissão Europeia a proposta de regulamento para estabelecer regras harmonizadas em matéria de Inteligência Artificial “*Artificial Intelligence Act*” (EUROPEAN UNION, 2021). Para a elaboração da proposta, foi aberta uma consulta pública, onde as partes interessadas concordaram majoritariamente numa abordagem regulatória baseada em diferentes três níveis de riscos associados à tecnologia¹⁷. Convertido em norma jurídica em 2024, o diploma traz consigo o artigo 50, o qual estabelece que provedores

¹⁷ Riscos Inaceitáveis – inclui sistemas de inteligência artificial que manipulem o comportamento humano e/ou explorem a vulnerabilidade de pessoas; sistemas de pontuação social; e, salvo exceções, identificação biométrica. Sistemas de Alto Risco – ao invés de trazer uma definição clara, a proposta classifica como de alto risco aqueles sistemas relacionados a uma lista de temas, discriminados nos anexos II e III. Entre eles estão aqueles já tutelados pela Legislação Harmonizada sobre Saúde e Segurança (tais como aparelhos médicos, brinquedos e máquinas) e aqueles relacionados a áreas especificadas (tais como treinamento educacional, emprego e aplicação da lei). A proposta impõe a obrigatoriedade de supervisão humana para esta categoria. Sistemas de Mínimo ou Baixo Risco – aqueles que não se enquadram nas categorias previstas acima.

devem garantir que sistemas de IA destinados a interagir com pessoas naturais sejam projetados e desenvolvidos de forma que aquelas sejam devidamente informadas de que estão interagindo com um sistema de IA, “salvo se tal for óbvio do ponto de vista de uma pessoa natural razoavelmente informada, atenta e advertida, tendo em conta as circunstâncias e o contexto de utilização”.

Estas obrigações devem recair sobre os sistemas independentemente de seus níveis de risco, visto que as exceções previstas se aplicam apenas para as hipóteses em que, pelas circunstâncias e contextos de uso, seja óbvio para o usuário que esteja se comunicando com uma IA, ou quando esta seja destinada a investigar e processar crimes.

Para os casos de sistemas não categorizados como de alto risco, não resta claro se a obrigação de transparência deveria alcançar o público externo nos casos em que o sistema de IA não é utilizado para interação com este, mas apenas para subsidiar ou determinar decisões internas de uma empresa, por exemplo, especialmente quando houver supervisão humana.

Para os provedores de sistemas da categoria de alto risco, a proposta impõe também a obrigação de registro junto a uma base de dados da União Europeia, cujas informações serão abertas para o público.

A regra geral, então, impõe a responsabilidade de *disclosure* aos provedores de sistemas, não aos usuários e às plataformas. Segundo a literatura, em certa medida, isso resolveria as objeções de especialistas que criticam a necessidade de identificação do usuário que utiliza o sistema automatizado, seja pessoa ou instituição (VEALE e BORGESIU, 2021). A doutrina aponta que a distinção entre provedor, usuário e autor pode, no entanto, colapsar na prática diante de aplicações como ChatGPT. Neste sentido, aponta, a título de exemplo, que uma eventual obrigação de que os textos produzidos pela ferramenta contivessem identificação (e.g. “#bot”) violaria a neutralidade tecnológica.

Cabe, por fim, destacar que as aplicações de sistemas automatizados também devem observar o regramento da Diretiva sobre Práticas Comerciais Desleais “*Unfair Commercial Practices Directive*” (EUROPEAN UNION, 2005), a qual proíbe ações que distorçam ou tendam a distorcer o comportamento econômico, quanto ao produto, tanto do consumidor médio quanto do consumidor médio de um grupo, quando a prática se dirige a uma determinada categoria de consumidores. A regra geral, que veda práticas comerciais consideradas desleais, também se aplica, portanto, àquelas empreendidas por meio de sistemas automatizados. Em outras palavras, estas ferramentas não podem ser utilizadas para práticas tendentes a induzir consumidores a tomarem decisões transacionais que não tomariam não fossem as informações falsas, enganosas, suscetíveis de induzir a erro, ou omitidas sobre o produto.

Nesse sentido, interessante ponto de vista é do Carissa Véliz, segundo a qual emojis não devem ser usados por *chatbots* (VÉLIZ, 2023), pois são mais eficientes em induzir reações, inclusive fisiológicas (endorfina e oxytoxina), sendo, portanto, poderosas ferramentas de manipulação; nesse sentido, poderiam ser compreendidos, inclusive, como *dark patterns* ou *deceptive patterns*¹⁸.

4 Panorama nacional: LGPD e Projeto de Marco Civil da IA

No direito brasileiro, as poucas normas sobre IA são setoriais¹⁹ (normas de conselhos profissionais e CNJ, por exemplo). Diante da ausência de uma norma geral, a IA continua em larga medida regulada por dispositivos que não foram pensados para regrá-la. Como resultado, uma considerável discussão tenta entender a IA pelo aspecto da responsabilidade. Aponta-se que a “IA não altera o sujeito passivo do dever de indenizar os danos sofridos pelo consumidor” (BESSA e NUNES, 2023, p. 13). No campo civil, pondera-se a imprevisibilidade do resultado e o ônus desproporcional que incide sobre o programador de IA, mas que tal imprevisibilidade também existiria nos casos de responsabilidade pelo fato da coisa e ato de terceiro (FONSECA, 2021, p. 19).

Ainda no direito posto, tem-se a Lei Geral de Proteção de Dados (LGPD) que se revelou sensível aos desafios da IA em seu âmbito material, razão inclusive para a elevada afinidade entre a LGPD e o projeto de Marco Civil da Internet (MACHADO *et al.* 2023, p. 2). Nesse espectro, caso se considere que a atividade da IA pode ser considerada utilização, processamento, avaliação ou extração de dados, a conclusão é que o titular tem o direito ao acesso facilitado às informações sobre o tratamento de seus dados, no art. 9º, LGPD. Mais do que isso, exceto nos casos de dados tornados manifestamente públicos pelo titular, isto dependerá do consentimento prévio (art. 7º, LGPD). Nesta hipótese, diante da vedação de tratamento de dados mediante vício de consentimento (art. 8º, §3º, LGPD) e do ônus probatório que recai sobre o controlador dos dados, chega a ser temerário omitir quando os dados serão objeto de tratamento automatizado por IA.

Ocorre que este cenário de risco jurídico ganha outros contornos sob a ótica do art. 20 da LGPD, o qual concede um direito subjetivo da pessoa de obter a revisão das decisões tomadas pela IA. Ausente limitação no texto legal, entende-se que tal direito se dá ainda que a decisão se prove completa e tecnicamente adequada. Não divulgar que a IA interagiu com o

¹⁸ Para uma melhor compreensão desses conceitos, vide especialmente: GDPR e Digital Services Act.

¹⁹ Vide, por exemplo, Recomendação n. 1/2024 – Conselho Federal da OAB; Resolução CNJ 615/2025; e Resolução CVM Nº 19/2021.

titular vulnera este direito à revisão. Portanto, o direito à identificação da IA decorre da garantia do direito (subjeto, faculdade) da pessoa de, querendo, ver revisada qualquer decisão tomada pela IA que atinja seus direitos.

Das várias redações de propostas de lei sobre o tema que tramitam no Congresso Nacional, escolhe-se para análise o Parecer de Plenário do Senado Federal quanto ao Projeto de Lei n.º 2.338/2023 (a seguir apenas “Parecer”), por ser atualmente o texto mais comentado na comunidade jurídica. No seu art. 5º, I²⁰, o Parecer dispõe sobre o direito de informação quanto à interação com sistema de IA, o que vai ao encontro da reflexão proposta neste artigo. Todavia, propõe-se quatro considerações sobre sua redação.

A primeira é compreender se o direito de informação deve ser exercido pelo titular ou se tal direito implica automaticamente em um dever do outro lado da relação jurídica. Em outras palavras, a identificação da IA, por exemplo, será espontânea ou só deve ocorrer se indagada pela pessoa que está interagindo? A segunda consideração toca à dúvida de quando a identificação deve ocorrer, se antes, durante ou depois da interação. Ao trazer o termo “afetado” (passado), a redação dá a entender que o direito à informação só surge após a pessoa ser “afetada”, isto é, tal direito não seria prévio à interação com a IA.

A terceira é relativa ao alcance do termo “afetado”. Ao prever o direito à informação independentemente do grau de risco da IA, a proposta parece ter expandido significativamente os casos em que seria exigido o direito à informação. Exercitando hipóteses, imagine-se uma IA operando uma prática de marketing olfativo (técnica que utiliza de aromas e fragrâncias para criar uma experiência olfativa atrativa ao potencial cliente de lojas físicas). A IA, então, customizaria os aromatizadores da loja, de forma sofisticada, de acordo com a temperatura, lotação da loja, entre outros fatores. Seria razoável aplicar o direito de informação aqui? Sendo um direito que implica em uma contrapartida automática (dever de informação), este hipotético

²⁰ Art. 5º A pessoa ou grupo afetado por sistema de IA, independentemente do seu grau de risco, tem os seguintes direitos, a serem exercidos na forma e nas condições descritas neste Capítulo:

I – direito à informação quanto às suas interações com sistemas de IA, de forma acessível, gratuita e de fácil compreensão, inclusive sobre caráter automatizado da interação, exceto nos casos em que se trate de sistemas de IA dedicados única e exclusivamente à cibersegurança e à ciberdefesa, conforme regulamento;

II – direito à privacidade e à proteção de dados pessoais, em especial os direitos dos titulares de dados nos termos da Lei nº 13.709, de 14 de agosto de 2018 (Lei Geral de Proteção de Dados Pessoais), e da legislação pertinente;

III – direito à não discriminação ilícita ou abusiva e à correção de vieses discriminatórios ilegais ou abusivos, sejam eles diretos ou indiretos.

§ 1º A informação referida no inciso I do caput deste artigo será fornecida com o uso de ícones ou símbolos uniformizados facilmente reconhecíveis, sem prejuízo de outros formatos.

§ 2º Os sistemas de IA que se destinem a grupos vulneráveis deverão, em todas as etapas de seu ciclo de vida, ser transparentes e adotar linguagem simples, clara e apropriada à idade e à capacidade cognitiva, e ser implementados considerando o melhor interesse desses grupos.

sistema de marketing olfativo deveria ser previamente identificado às pessoas “afetadas” à medida que operasse. Não parece ter sido esta a intenção dos proponentes do texto do Parecer.

A quarta e última reflexão está atrelada à terceira: o direito à informação existe para todo e qualquer tipo de interação com a IA, mesmo que indireta e passiva? Como harmonizar o previsto no mencionado art. 5º com o tratado no art. 19²¹ do Parecer (“conteúdo sintético”)? Assim como em relação à terceira reflexão, sugere-se avaliar uma limitação a esse direito de informação para não gerar um dever desproporcional. A melhor interpretação de “afetado” e do tipo de interação que concede o direito à informação seria, então, restrita às hipóteses em que se extingue, cria-se ou modifica-se direitos?

5 Um Captcha²² inverso é realmente necessário?

A IA trouxe um cipoal de controvérsias, nas mais variadas esferas jurídicas, desde o questionamento acerca da segurança do mercado financeiro à solidez do regime democrático, da revisitação às bases do direito da personalidade aos institutos de direito societário, da releitura do direito concorrencial e às novas fronteiras da propriedade intelectual.

No presente artigo, buscou-se, contudo, realizar uma avaliação panorâmica do dever de transparência atribuído à IA, focando mais nas regras e menos nas exceções. Assim, ficam ressalvadas aquelas hipóteses relativas a interesse público, a exemplo de pessoas vulneráveis, com regramento por microsistemas, tais como Estatuto do Idoso e Estatuto da Criança e do Adolescente.

A assimetria de informação é usualmente regulada em relações não paritárias ou quando, de alguma forma, efetivamente representar um risco. No Parecer avaliado, chama atenção a previsão do direito à informação independentemente da classificação de risco da IA e sem maiores detalhes de quando esse direito deve ser tutelado.

²¹ Art. 19. Quando o sistema de IA gerar conteúdo sintético, deverá, considerando o estado da arte do desenvolvimento tecnológico e o contexto de uso, incluir identificador em tais conteúdos para verificação de autenticidade ou de características de sua proveniência, modificações ou transmissão, conforme regulamento.

§ 1º A presença do identificador previsto no caput não supre outros requisitos de informação e transparência, bem como outros parâmetros a serem definidos em regulamento.

§ 2º A autoridade competente, em colaboração com o Conselho Permanente de Cooperação Regulatória de Inteligência Artificial (Cria), disponibilizará biblioteca de softwares com vistas a facilitar o cumprimento da obrigação de sinalização, idealmente adotando padrão internacional amplamente reconhecido.

§ 3º O uso de conteúdo sintético em obras com finalidade artística, cultural ou de entretenimento poderá, sempre que não representar risco de disseminação de informações falsas, ser sinalizado por meios que não comprometam a utilidade e a qualidade da obra, como nos créditos ou nos metadados associados a tal obra, preservando sua fruição pelo público e seus usos convencionais.

²² Sigla de *Completely Automated Public Turing test to tell Computers and Humans Apart*, mecanismo de segurança popularmente utilizado no ambiente digital para distinguir usuários humanos e robôs (programas automatizados).

A iniciativa legislativa brasileira está em harmonia com a tendência internacional, a qual deve ser considerada, afinal a tecnologia é transfronteiriça; muito além, é, aliás, questão geopolítica atualmente. Contudo, nos termos suscitados neste artigo, especialmente no tópico anterior, cabe ainda um aprimoramento da previsão relativa à identificação da IA (direito de informação), de modo que não se configure excesso de regulação.

Referências bibliográficas

ALBUQUERQUE, K. TikTok trabalha em selo para identificar conteúdo gerado por IA. Tecmundo. S.l. Disponível em: <<https://www.tecmundo.com.br/software/267564-tiktok-trabalha-selo-identificar-conteudo-gerado-ia.htm?ab=true>>. Acesso em 14/03/2025.

ANTAELLA, Lucia. Pós-humano: por quê? In: *Revista USP*, n. 74, p. 126-137, jun./ago. 2007. Disponível em: <https://www.revistas.usp.br/revusp/article/view/13607>. Acesso em: 17 jun. 2018.

BASTOS, P. Neil ouve cores e é hackeável. É o primeiro ciborgue oficial do mundo. CNN Portugal, S.l, 2023. Disponível em: <<https://cnnportugal.iol.pt/videos/neil-ouve-cores-e-e-hackeavel-e-o-primeiro-ciborgue-oficial-do-mundo/6442ded0cf2665294e02821>>. Acesso em 14/03/2025.

BBC. A imagem feita por inteligência artificial que enganou jurados de um grande prêmio de fotografia. G1. São Paulo: 2023, Disponível em:< <https://g1.globo.com/pop-arte/noticia/2023/04/18/a-imagem-feita-por-inteligencia-artificial-que-enganou-jurados-de-um-grande-premio-de-fotografia.ghtml>>. Acesso em: 14/03/2025.

BENJAMIN, A. H. V., MARQUES, C. L., BESSA, L. R. *Manual de direito do consumidor*, 10ª ed. São Paulo: Revista dos Tribunais, 2022. Edição eletrônica.

BOLZANI, I. CONAR decide arquivar processo contra propaganda que recriou Elis Regina com inteligência artificial. G. São Paulo, 2023. Disponível em: <https://g1.globo.com/economia/midia-e-marketing/noticia/2023/08/23/conar-decide-arquivar-processo-contr-propaganda-que-recriou-elis-regina-com-inteligencia-artificial.ghtml>. Acesso em: 14/03/2025.

CALIFORNIA. SB-1001 Bots: disclosure. [S. l.], 2017. Disponível em: https://leginfo.legislature.ca.gov/faces/billStatusClient.xhtml?bill_id=201720180SB1001. Acesso em: 14 mar. 2025.

COMISSÃO EUROPÉIA. Document 52018dc0236: Communication from the Commission to the European Parliament, The Council, The European Economic and Social Committee and the Committee of the Regions Tackling online disinformation: a European Approach. [S. l.], 2018. Disponível em: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52018DC0236>. Acesso em: 14 mar. 2025.

DAVALOS, J. LANXON, N. Antony Levandowski Reboots Church of Artificial Intelligence. Bloomberg. S. L. 2023. Disponível em: <<https://www.bloomberg.com/news/articles/2023-11-23/anthony-levandowski-reboots-the-church-of-artificial-intelligence?srnd=technology-vp>>. Acesso em 14/03/2025.

EUROPEAN UNION. EU Code of Practice on Disinformation. [S. l.], 2018. Disponível em: <https://ec.europa.eu/newsroom/dae/redirection/document/87534>. Acesso em: 14 mar. 2025.

USP. Dilemas da inteligência artificial responsável. São Paulo, 2024. Disponível em: <<https://www.scoop.it/topic/educacao-3-0-uma-jornada?q=An%C3%BAncios%20que%20inclu%C3%ADram%20r%C3%B3tulos%20sobre%20o%20uso%20de%20IA%20foram%20classificados%20pelos%20entrevistados%20como%20menos%20confi%C3%A1veis%2C%20menos%20interessantes%20e%20menos%20>>. Acesso em 14/03/2025.

EUROPEAN UNION. THE Strengthened Code of Practice on Disinformation 2022. [S. l.], 2022. Disponível em: <https://ec.europa.eu/newsroom/dae/redirection/document/87585>. Acesso em: 14 mar. 2025.

EUROPEAN UNION. Digital Services Act. [S. l.], 2022. Disponível em: <<https://ec.europa.eu/newsroom/dae/redirection/document/87585>>. Acesso em: 14 mar. 2025.

EUROPEAN UNION. *Artificial Intelligence Act*. [S. l.], 2021. Disponível em: <<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206>>. Acesso em: 14 mar. 2025.

EUROPEAN UNION, Document 32005L0029: Directive 2005/29/EC of the European Parliament and of the Council of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market and amending Council Directive 84/450/EEC, Directives 97/7/EC, 98/27/EC and 2002/65/EC of the European Parliament and of the Council and Regulation (EC) No 2006/2004 of the European Parliament and of the Council (‘Unfair Commercial Practices Directive’) (Text with EEA relevance). [S. l.], 2005. Disponível em: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32005L0029>. Acesso em: 14 mar. 2025.

FERREIRA, T. OpenAI e outras terão selo para identificar conteúdo gerado por IA. Olhar digital. S.l. 2023. Disponível em: <<https://olhardigital.com.br/2023/07/21/internet-e-redes-sociais/openai-e-outras-terao-selo-para-identificar-conteudo-gerado-por-ia/>>. Acesso em: 14/03/2025.

GARCIA, S. C. Revisão dos Contratos. In: *Doutrinas Essenciais Obrigações e Contratos*, vol. 4. São Paulo: Revista dos Tribunais, 2011.

GRANT, C. A Manipulação Discursiva e a Figura do “Homem Médio” no Direito Penal. In: *Ciências Penais*, vol. 9. São Paulo: Revista dos Tribunais, 2008. *Texto online*.

HARAWAY, Donna Jeanne. Manifesto ciborgue: ciência, tecnologia e feminismo-socialista no final do século XX. In: TADEU, Tomaz. (Org.). *Antropologia do ciborgue: as vertigens do pós-humano*. Trad. Tomaz Tadeu. 2. ed. Belo Horizonte: Autêntica, 2009.

KUMAR, S. (ed). Circuits: Newsletter of the Computer & Technology Section of the State Bar of Texas. [S.l.], 2019. Disponível em <<https://sbot.org/wp-content/uploads/2020/05/Circuits-December2019.pdf>>. Acesso em 14/03/2025.

KURZWEIL, Ray. The singularity is near. In: *Ethics and Emerging Technologies*. Palgrave Macmillan, London, 2014.

AMO, M. CALO, R. Regulating Bot Speech, 66 UCLA L. Rev. 988 (2019), Disponível em: <<https://digitalcommons.law.uw.edu/faculty-articles/532>>.

LUO, X. TONG, S. FANG, Z. e QU, Z. “Frontiers: Machines vs. Humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases”. In: *Informing Marketing Science*. Sep., 2019. Disponível em <<https://pubsonline.informs.org/doi/abs/10.1287/mksc.2019.1192>>. Acesso em 14/03/2025.

MCCALLUM, S. ChatGPT banned in Italy over privacy concerns. BBC. Londres, 2023. Disponível em: < <https://www.bbc.com/news/technology-65139406> >. Acesso em 14/03/2025.

THE MICROSOFT Responsible AI Standard: Explore Microsoft internal guidance on how to design, build, and test AI systems. [S. l.], 2019. Disponível em: <https://www.microsoft.com/en-us/ai/principles-and-approach#coreui-areaheading-1ordjnc>. Acesso em: 14 mar. 2025. MIGALHAS, Conar arquiva processo contra Volks por comercial com Elis Regina. S.l. 2023. Disponível em: <<https://www.migalhas.com.br/quentes/392385/conar-arquiva-processo-contra-volks-por-comercial-com-elis-regina>>.

MILLER, Z. NICKS, D. Meet the Robot Telemarketer Who Denies She's A Robot. Time. S. l., 2013. Disponível em: <<https://newsfeed.time.com/2013/12/10/meet-the-robot-telemarketer-who-denies-shes-a-robot/>>. Acesso em: 14 mar. 2025.

NASCIMENTO, A. Matrimônio 2.0: o casamento entre humanos e robôs. Estadão. São Paulo, 2022. Disponível em: < https://www.estadao.com.br/link/alexandre-nascimento/matrimonio-2-0/?srsltid=AfmBOoqQKOQ1B3zYLQnRfOVdrUpBZPB1C3Uz_5_BTl1uwCIhwTBiMY51 >. Acesso em 14/03/2025.

NICKS, D. Robot Telemarketer Employer: Samantha West Is No Robot. S. l., 2013. Disponível em: <<https://newsfeed.time.com/2013/12/17/robot-telemarketer-samantha-west/>>. Acesso em: 14 mar. 2025.

O GLOBO, Meta identificará em suas redes sociais áudios, imagens e vídeos gerados por IA. S. l., 2024. Disponível em: < <https://oglobo.globo.com/economia/tecnologia/noticia/2024/04/05/meta-identificara-em-suas-redes-sociais-audios-imagens-e-videos-gerados-por-ia.ghtml> >. Acesso em 14/03/2025.

OPEN LOOP. Automated Decision Impact Assessment (ADIA). S. l., 2020. Disponível em: <https://openloop.org/programs/ai-impact-assessment-europe/>. Acesso em: 14/03/2025.

PEARLMAN, Ellen. I. Cyborg. In: *Performing Arts Journal*, n. 110, p. 84-110, 2015. Disponível em: [www.academia.edu/12592330/I_Cyborg]. Acesso em: 17/06/2018.

RAMOS, R. Mulher se casa com robô de inteligência artificial. BAND. São Paulo: 2023. Disponível em <<https://www.band.uol.com.br/band-fm/noticias/mulher-se-casa-com-robo-de-inteligencia-artificial-16607948>>. Acesso em: 14/03/2025.

RAPOPORT, D., “The Four Waves of Modern Terrorism”. In: Audrey Kurth Cronin and James M. Ludes, *Attacking Terrorism: Elements of a Grand Strategy*, Washington: Georgetown University Press, 2004, p. 46-73.

RIVELLI, F. Do Blueprint à execução: Uma análise da estratégia americana para governança da Inteligência Artificial. Migalhas, S.l. <<https://www.migalhas.com.br/coluna/ia-em-movimento/422435/do-blueprint-a-execucao-estrategia-americana-para-governanca-da-ia>>. Acesso em: 14/03/2025.

SAVAGE, M. Por que música viral feita por inteligência artificial com vozes de Drake e The Weeknd preocupa artistas. BBC. Londres, 2023. Disponível em:< <https://www.bbc.com/portuguese/articles/cy0r0zny1pko> >. Acesso em: 14/03/2025.

UNITED STATES CONGRESS, H.R.4536 - Bot Disclosure and Accountability Act of 2019. [S. l.], 2019. Disponível em: <https://www.congress.gov/bill/116th-congress/house-bill/4536>. Acesso em: 14 mar. 2025.

UNITED STATES CONGRESS. S.2125 - Bot Disclosure and Accountability Act of 2019. [S. l.], 2019. Disponível em: <https://www.congress.gov/bill/116th-congress/senate-bill/2125#:~:text=This%20bill%20requires%20social%20media,the%20provider's%20social%20media%20website..> Acesso em: 14/03/2025.

UNITED STATES CONGRESS. S.3127 A bill to protect the right of the American public under the First Amendment to the Constitution of the United States to receive news and information from disparate sources by regulating the use of automated software programs intended to impersonate or replicate human activity on social media. [S. l.], 2018. Disponível em: <https://www.congress.gov/bill/115th-congress/senate-bill/3127>. Acesso em: 14/03/2025.

SENTINEL. Defending Against Deepfakes and Information Warfare. [S. l.], 2020. Disponível em: 2020. Acesso em: 14/03/2025.

STRICKLIN, K. Social Media Bots: Laws, Regulations, and Platform Policies. CNA. Washington, 2020. Disponível em: < <https://apps.dtic.mil/sti/trecms/pdf/AD1112566.pdf> >. Acesso em 14/03/2025.

UNITED STATES, Constitution of the United States, first amendment. Washington, [s.d]. Disponível em: <<https://constitution.congress.gov/constitution/amendment-1/>>. Acesso em: 14/03/2025.

UNITED STATES SUPREME COURT, *Talley vs. California*, 362 U.S. 60 (1960). Disponível em: <<https://supreme.justia.com/cases/federal/us/362/60/>>. Acesso em: 14/03/2025.

UNITED STATES SUPREME COURT, *McIntyre v. Ohio Elections Commission*, 3 514 U.S. 334 (1995). Disponível em: < <https://supreme.justia.com/cases/federal/us/514/334/>>. Acesso em: 14/03/2025.

VARGAS, E. Elon Musk, Yuval Harari e mais de mil especialistas pedem ‘pausa da Inteligência Artificial’ em carta. Suno. São Paulo, 2023. Disponível em: < <https://www.suno.com.br/noticias/inteligencia-artificial-musk-harari-carta/>>. Acesso em: 14/03/2025.

VÉLIZ, C. Chatbots shouldn’t use emojis. Nature. Berlin, 2023. Disponível em: <<https://philpapers.org/archive/VLICSU.pdf> >. Acesso em 14/03/2025.

VEALE, M. BORGESIU, F. Z. Demystifying the Draft EU Artificial Intelligence Act. In: Computer Law Review International. Berlin, DeGrutyer, 2021. Disponível em: < <https://www.degruyter.com/document/doi/10.9785/cr-2021-220402/html>>. Acesso em 14/03/2025.

THE WHITE HOUSE. Blueprint for an AI Bill of Rights. Washington. Disponível em: <https://bidenwhitehouse.archives.gov/ostp/ai-bill-of-rights/>. Acesso em: 14/03/2025.

THE WHITE HOUSE. What is the Blueprint for an AI Bill of Rights?. Washington. Disponível em: <https://bidenwhitehouse.archives.gov/ostp/ai-bill-of-rights/what-is-the-blueprint-for-an-ai-bill-of-rights/>. Acesso em: 14/03/2025.