

Economic Analysis of Law Review

A Inteligência Artificial a Serviço da Desinformação: como as *Deepfakes* e as Redes Automatizadas Abalam a Liberdade de Ideias no Debate Público e a Democracia Constitucional e Deliberativa

Artificial Intelligence in the Service of Disinformation: How Deepfakes and Automated Networks Attack the Freedom of Ideas in the Public Debate and Constitutional and Deliberative Democracy

Ilton Norberto Robl Filho¹
Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa
(IDP)

Marco Aurélio Marrafon²
Universidade do Estado do Rio de Janeiro (UERJ)

Filipe Medón³
Pontifícia Universidade Católica do Rio de Janeiro (PUC-
RJ)

RESUMO

A internet e as redes sociais modificaram profundamente as relações comunicativas e a esfera pública, reduzindo as formas de interação pluralista e criando “bolhas” de discussão, em que seus membros possuem pensamentos, reflexões e propostas muito homogêneas. Por sua vez, o artigo analisa as *deepfakes* e as redes automatizadas, as quais ampliam os dilemas e os problemas na democracia constitucional e deliberativa, que pressupõe igualdade, liberdade e pluralismo nos espaços comunicativos. Com exceção do emprego para humor das *deepfakes*, observa-se a violação da liberdade de expressão e do direito à informação, pois informações falsas são disseminadas com aparência de verossimilhança para receptores menos atentos e o livre mercado de ideias torna-se ainda mais desigual. O artigo conclui pela necessidade de controle sobre as *deepfakes* e das redes automatizadas deve focar na constatação de padrões inautênticos de atuação, evitando-se na maior medida possível a análise dos conteúdos para evitar atos de censura ainda que bem-intencionados e sendo o controle necessário especialmente no contexto de populismo digital e novas democraturas. O artigo dialoga com a literatura especializada, nas searas do direito, ciência política, filosofia política e tecnologia, utilizando assim a revisão bibliográfica como instrumento metodológico e refletindo sobre as realidades social e econômica.

Palavras-chave: Democracia. Constituição. Liberdade Comunicacional. *Deepfakes*. Inteligência Artificial.

JEL: K14, K42.

ABSTRACT

The internet and social networks have profoundly modified the communicative relations and the public sphere, reducing the forms of pluralist interaction and creating discussion “bubbles”, in which its members have very homogeneous thoughts, reflections and proposals. In turn, the article analyzes deepfakes and automated networks, which amplify the dilemmas and problems in constitutional and deliberative democracy, which presupposes equality, freedom and pluralism in communicative spaces. With the exception of the use of deepfakes for humor, there is a violation of freedom of expression and the right to information, as false information is disseminated with an appearance of verisimilitude to less attentive recipients and the free market of ideas becomes even more unequal. The article concludes that there is a need for control over deepfakes and automated networks should focus on the verification of inauthentic patterns of performance, preventing as much as possible the analysis of the contents to avoid acts of censorship, even if well intentioned, and the control being necessary especially in the context of digital populism and new democratures. The article dialogues with the specialized literature, in the fields of law, political science, political philosophy and technology, thus using the literature review as a methodological instrument and reflecting on social and economic realities.

Keywords: Democracy. Constitution. Communicational Freedom. Deepfakes. Artificial intelligence.

R: 20/10/20 **A:** 22/02/22 **P:** 31/12/22

¹ E-mail: norbertorobl@gmail.com.

² E-mail: marco_marrafon@yahoo.com.br.

³ E-mail: filipemedon@hotmail.com

1. Introdução

“Uma mentira pode viajar meio mundo enquanto a verdade está calçando os sapatos”, já dizia Mark Twain. Ou melhor: Winston Churchill. Aliás, dizia o satirista Jonathan Swift. Ou, talvez, nenhum deles. É curioso notar que a frase tantas vezes repetida quando se fala do fenômeno da desinformação seja, ela própria, carregada de imprecisões quanto ao seu verdadeiro autor, como denuncia reportagem do jornal *The New York Times*.⁴ No entanto, seja quem for o seu emissor original, a frase, que teria pelo menos um século de existência e já viajou o mundo amarrada aos sapatos de diversos autores, parece mais atual do que nunca, especialmente em face da potencialização no emprego de técnicas de Inteligência Artificial para desinformar.

Em julho de 2020, em meio a discussões sobre a aprovação do chamado Projeto de Lei das Fake News (PL n.º. 2630/2020) e de suas controvérsias, o Facebook tomou duas ações contundentes para dismantelar redes de desinformação no Brasil, atingindo dois lados da polarizada política brasileira: apoiadores do Presidente Jair Bolsonaro e, de outro lado, do Partido dos Trabalhadores (PT).

As ações buscaram combater, no primeiro caso, uma rede de contas falsas relacionadas ao Partido Social Liberal (PSL) e a gabinetes da família Bolsonaro. Segundo reportagem do G1, o “grupo publicava conteúdo sobre notícias, eventos, eleições, memes, críticas à oposição e a jornalistas.”⁵ De acordo com o Facebook, os responsáveis tentavam ocultar identidade, ou seja, “essas contas estavam envolvidas com a criação de perfis falsos e com ‘comportamento inautêntico’ — quando um grupo de páginas e pessoas atuam em conjunto para enganar outros usuários sobre quem são e o que estão fazendo.”⁶ A plataforma ainda declarou que, quando investiga e remove esse tipo de operação, concentra-se mais “no comportamento, e não no conteúdo — independentemente de quem esteja por trás dessas redes, qual conteúdo elas compartilhem, ou se elas são estrangeiras ou domésticas.”⁷

Por seu turno, “[o] PT teve nove de suas contas no WhatsApp desativadas nas duas últimas semanas por envio automatizado de mensagens, em um processo que envolveu denúncias de spam político. As duas ações são proibidas pelos termos de uso da plataforma.”⁸ As atividades dos canais do PT foram detectadas por sistemas de aprendizado de máquina (*machine learning*) usados pelo próprio WhatsApp, que, por exemplo, identifica que um envio de mais de uma centena de mensagens por minuto seria incompatível com uma atuação humana.⁹ Como se pode

⁴ CHOKSHI, Niraj. That Wasn't Mark Twain: How a Misquotation Is Born. In: *The New York Times*, 26 abr. 2017. Disponível em: <<https://www.nytimes.com/2017/04/26/books/famous-misquotations.html>> Acesso em 08 jul. 2020.

⁵ G1. *Facebook remove rede de contas falsas relacionada ao PSL e a gabinetes da família Bolsonaro*. 08 jul. 2020. Disponível em: <<https://g1.globo.com/economia/tecnologia/noticia/2020/07/08/facebook-remove-rede-de-contas-falsas-relacionada-ao-psl-e-a-gabinetes-da-familia-bolsonaro.ghtml>> Acesso em 08 jul. 2020.

⁶ G1. *Facebook remove rede de contas falsas relacionada ao PSL e a gabinetes da família Bolsonaro*. 08 jul. 2020. Disponível em: <<https://g1.globo.com/economia/tecnologia/noticia/2020/07/08/facebook-remove-rede-de-contas-falsas-relacionada-ao-psl-e-a-gabinetes-da-familia-bolsonaro.ghtml>> Acesso em 08 jul. 2020.

⁷ G1. *Facebook remove rede de contas falsas relacionada ao PSL e a gabinetes da família Bolsonaro*. 08 jul. 2020. Disponível em: <<https://g1.globo.com/economia/tecnologia/noticia/2020/07/08/facebook-remove-rede-de-contas-falsas-relacionada-ao-psl-e-a-gabinetes-da-familia-bolsonaro.ghtml>> Acesso em 08 jul. 2020.

⁸ GOMES, Helton Simões. WhatsApp barra contas do PT por disparos em massa; partido vê "interesse". In: *Tilt*, 07 jul. 2020. Disponível em: <https://www.uol.com.br/tilt/noticias/redacao/2020/07/07/whatsapp-bloqueia-contas-do-pt-flagrado-por-disparo-em-massa-de-mensagens.htm?aff_source=56d95533a8284936a374e3a6da3d7996> Acesso em 09 jul. 2020.

⁹ GOMES, Helton Simões. WhatsApp barra contas do PT por disparos em massa; partido vê "interesse". In: *Tilt*, 07 jul. 2020. Disponível em: <https://www.uol.com.br/tilt/noticias/redacao/2020/07/07/whatsapp-bloqueia-contas-do-pt-flagrado-por-disparo-em-massa-de-mensagens.htm?aff_source=56d95533a8284936a374e3a6da3d7996> Acesso em 09 jul. 2020.

A Inteligência Artificial a Serviço da Desinformação: como as *Deepfakes* e as Redes Automatizadas Abalam a Liberdade de Ideias no Debate Público e a Democracia Constitucional e Deliberativa

notar, a plataforma recorreu à própria tecnologia da Inteligência Artificial para coibir atos por ela reputados como ilícitos perpetrados pelo mesmo recurso tecnológico.

Para além das redes automatizadas de envio em massa de mensagens, a Inteligência Artificial (IA), também vem sendo empregada para desinformar por meio das chamadas *deepfakes*, isto é, reconstruções digitais da imagem de uma pessoa, que passa a ser retratada de forma falsa, normalmente com o intuito de provocar desinformação, embora também esteja sendo amplamente utilizada para o humor e a sátira.¹⁰ Veja-se, por exemplo, o caso da *deepfake* do atual governador de São Paulo, João Dória, que às vésperas das eleições do ano de 2018 foi vinculado a um vídeo em que supostamente o então candidato participava de uma orgia com mulheres. Peritos teriam constatado se tratar de uma montagem, isto é, uma *deepfake*,¹¹ criada para abalar a imagem e a honra do governador, com influência direta na acirrada disputa eleitoral.

Nesse contexto político, recentemente até mesmo o presidente dos EUA Donald Trump compartilhou um vídeo de um discurso da presidente da Câmara dos Deputados, Nancy Pelosi, em velocidade levemente desacelerada, que, embora não se tratasse de uma *deepfake* no sentido original, serve como alerta para os riscos do emprego das técnicas de Inteligência Artificial para finalidades antidemocráticas.¹²

É importante advertir desde logo que, embora a desinformação não esteja necessariamente atrelada ao cenário eleitoral, sua utilização mais frequente e notória tem se dado para esta finalidade, podendo ter impactos sensíveis para o jogo democrático, sobretudo em quadros de intensa polarização ideológica e pouca margem de vantagem entre os principais candidatos. É o que se viu em recentes sufrágios, como nas eleições norte-americanas de 2016, no Brexit, que culminou na saída do Reino Unido da União Europeia, e nas eleições brasileiras de 2018.

Considerado esse cenário, o presente artigo busca analisar o impacto das técnicas de Inteligência Artificial aplicadas no *deepfake* e nas redes automatizadas para o fenômeno da desinformação, que, como se verá, pode ter consequências catastróficas para a democracia. É preciso, no entanto, ressaltar a premissa fundante de que a tecnologia não pode ser demonizada, pois “[n]ossas interações com robôs e sistemas de IA são interações com as pessoas que estão implementando essas novas tecnologias, mesmo quando nós não o percebemos”.¹³ Dito diversamente, “é preciso saber que há poderes invisíveis que buscam se ocultar por trás da tecnologia”^{14 15} e que o alvo das críticas e censuras deve recair sobre esses sujeitos elípticos.

¹⁰ Recomenda-se, a esse respeito, as vídeo-montagens realizadas por Brunno Sartori no Instagram.

¹¹ FOLHA DE SÃO PAULO. Peritos constataram montagem em vídeo vazado, afirma Dória. 24 out. 2018. Disponível em: <<https://www1.folha.uol.com.br/poder/2018/10/peritos-constataram-montagem-em-video-vazado-afirma-doria.shtml>> Acesso em 09 jul. 2020.

¹² AFFONSO, Carlos. Tecnologia abre novo capítulo na manipulação de vídeos; imagina na eleição. In: *Tecfront*, 28 mai. 2019. Disponível em: <<https://tecfront.blogosfera.uol.com.br/2019/05/28/tecnologia-abre-novo-capitulo-na-manipulacao-de-vidEOS-imagina-na-eleicao/>> Acesso em 10 abr. 2020.

¹³ No original: “Our interactions with robots and AI systems are interactions with the people who are deploying these new technologies, even when we do not realize it.” (BALKIN, Jack. The path of robotics law. *California Law Review Circuit*, Berkeley, v. 06, p. 59, jun. 2015).

¹⁴ MEDON, Filipe. *Inteligência Artificial e Responsabilidade Civil: autonomia, riscos e solidariedade*. Salvador: JusPodivm, 2020, p. 178.

¹⁵ No mesmo sentido: “First, if we are to understand how technology impacts society in general, and politics and democratic communications in particular, we must not be caught up in the particular, novel, technical disruption. Instead, we have to expand our viewpoint across time and systems, and understand the long term structural interactions between technology, institutions, and culture. Through this broader and longer- term lens, the present epistemic crisis is not made of technology; it cannot be placed at the feet of the internet, social media, or artificial intelligence. It is a phenomenon rooted in the radicalization of the right wing of American politics and a thirty-year

Na seção 2 serão analisadas as *deepfakes*, e as redes automatizadas de desinformação e a inteligência artificial. Por sua vez, a seção 3 reflete sobre a democracia constitucional deliberativa, as liberdades comunicacionais e o direito à informação nesse novo contexto de debate eletrônico e digital. Ao final, são apresentadas as principais conclusões.

2. Entre *Deepfakes* e Redes Automatizadas de Desinformação

O fenômeno da desinformação se popularizou em 2016 com o nome de *fake news*¹⁶, ou notícias falsas, as quais, na definição de Carlos Affonso Souza e Chiara de Teffé, “podem ser compreendidas como conteúdos falsos, inverídicos, distorcidos ou fora de contexto que são espalhados, como se notícias fossem, para promover propositalmente desinformação ao público”.¹⁷ Ou seja, destaca-se, de imediato, o caráter intencional da prática, pelo que já se diferencia, por exemplo, do erro jornalístico, isto é, a notícia que, ainda que por culpa de seu autor, divulga fato em desacordo com a verdade. A desinformação não se confunde, ademais, com o humor e a sátira, que, proposital e claramente, mentem para divertir e sugerir a reflexão sobre temas sensíveis.

E a tecnologia, em especial a Inteligência Artificial, tem sido utilizada para potencializar a propagação de conteúdos falsos com esse propósito deliberado de desinformar. Yochai Benkler, Robert Faris e Hal Roberts esclarecem como constataram esta atuação no cenário norte-americano:

Algo fundamental estava acontecendo para ameaçar a democracia, e nosso olhar coletivo pousou na nova e rápida mudança - a tecnologia. Processos tecnológicos fora do controle de qualquer pessoa ou país - a convergência de mídia social, curadoria algorítmica de notícias, bots, inteligência artificial e análise de big data - estavam criando câmaras de eco que reforçavam nossos preconceitos, removiam indícios de confiabilidade e eram geralmente opressores sobre nossa capacidade de entender o mundo e, com ela, nossa capacidade de autogoverno como democracias razoáveis.¹⁸

process of media markets rewarding right-wing propagandists. We suspect that a similarly broad and long-term lens will be required to properly understand the rise of far-right parties and their information ecosystems elsewhere. At least in the United States, we find here that a failure to do so results in severe misdiagnosis of the challenges we face.” (BENKLER, Yochai; FARIS, Robert; ROBERTS, Hal. *Network propaganda: manipulation, disinformation, and radicalization in American politics*. New York: Oxford University Press, 2018, p. 42).

¹⁶ “The term fake news was initially used to describe falsehoods that were propounded and given wide circulation on the internet. Now even the term fake news itself has been drained of meaning, used as a way to describe anything the speaker disagrees with. Although some social media platforms have taken steps to counter it, the nature of online communication (as currently engineered) is conducive to the rapid spread of misinformation. The result is so-called post-truth politics.” (SUSSKIND, Jamie. *Future Politics: living together in a world transformed by tech*. Oxford: Oxford University Press, 2018, p. 230).

¹⁷ SOUZA, Carlos Affonso Pereira de; TEFFÉ, Chiara Spadaccini de. *Fake news: como garantir liberdades e conter notícias falsas na internet?* In: MENEZES, Joyceane Bezerra de; TEPEDINO, Gustavo (Coord.), *Autonomia Privada, liberdade existencial e direitos fundamentais*. Belo Horizonte: Fórum, 2019, pp. 525-543.

¹⁸ Tradução livre de: “Something fundamental was happening to threaten democracy, and our collective eye fell on the novel and rapidly changing — technology. Technological processes beyond the control of any person or country — the convergence of social media, algorithmic news curation, bots, artificial intelligence, and big data analysis— were creating echo chambers that reinforced our biases, were removing indicia of trustworthiness, and were generally overwhelming our capacity to make sense of the world, and with it our capacity to govern ourselves as reasonable democracies” (BENKLER, Yochai; FARIS, Robert; ROBERTS, Hal. *Network propaganda: manipulation, disinformation, and radicalization in American politics*. New York: Oxford University Press, 2018, p. 4).

A Inteligência Artificial a Serviço da Desinformação: como as *Deepfakes* e as Redes Automatizadas Abalam a Liberdade de Ideias no Debate Público e a Democracia Constitucional e Deliberativa

Ainda naquele país, questionou-se acerca da coleta de dados para o direcionamento de propaganda política por meio de manipulação comportamental, que se valia de técnicas de Inteligência Artificial para analisar perfis de eleitores.¹⁹

Ao lado dessa criação de perfis a partir de dados coletados dos eleitores para direcionamento e bombardeamento de conteúdo, seja ele falso ou não, destacam-se as redes automatizadas e as *deepfakes*, como se demonstrará a seguir.

Essas redes acabam gerando o chamado “comportamento inautêntico coordenado” (*Coordinated Inauthentic Behavior*, ou CIB). Nesse sentido, Nathaniel Gleicher, diretor de políticas de segurança do Facebook, “já afirmou que uma peça central para identificar comportamento inautêntico coordenado na plataforma é o uso de contas falsas, duplicadas, que servem para aumentar a audiência umas das outras, enganando o público”.²⁰

Após a organização de dados pessoais obtidos de diversos modos (até mesmo ilegalmente fornecidos por empresas de cobrança ou por funcionários de empresas de telefonia), empresas dedicadas a esse tipo de atividade fazem disparos em massa se valendo de algoritmos que “segmentam os membros dos grupos entre apoiadores, detratores e neutros, e, desta maneira, conseguem customizar de forma mais eficiente o tipo de conteúdo que enviam.”²¹ Segundo Rodrigo Jorge, especialista em segurança da informação, “[o] que torna essas empresas eficientes é que elas criam robôs que simulam uma ação no aplicativo, como se mandassem mensagens individuais. Elas não são encaminhadas em listas de transmissão.”²² Estes “robôs usados pelas agências ainda conseguem saber se um número da lista de contatos informada pelo cliente possui ou não conta no WhatsApp.”²³ Tudo isso financiado, em muitos casos, de maneira não contabilizada pela Justiça Eleitoral como gasto declarado de campanha.

Há, além desses, os chamados “robôs sociais” (*social bots*), que atuam dentro de redes de robôs (*botnets*) e que “são contas controladas por *software* que geram artificialmente conteúdo e estabelecem interações com não robôs. Eles buscam imitar o comportamento humano e se passar como tal de maneira a interferir em debates espontâneos e criar discussões forjadas.”²⁴ Pautam, assim, as discussões nas redes sociais, ainda que não disseminem conteúdos falsos, porque podem simplesmente massificar o debate sobre um tema irrelevante, afastando o foco de um assunto delicado ou até mesmo de uma notícia contrária a determinado político ou pessoa que estaria, se não houvesse a interferência artificial, na pauta do dia.

¹⁹ BENKLER, Yochai; FARIS, Robert; ROBERTS, Hal. *Network propaganda: manipulation, disinformation, and radicalization in American politics*. New York: Oxford University Press, 2018, p. 11.

²⁰ AFFONSO, Carlos. O que difere a ação de perfis bolsonaristas no Facebook e de fãs do K-pop? *Tecfront*, 10 jul. 2020. Disponível em: <<https://tecfront.blogosfera.uol.com.br/2020/07/10/o-que-difere-a-acao-de-perfis-bolsonaristas-no-facebook-e-de-fas-do-k-pop/>> Acesso em 10 jul. 2020.

²¹ MELLO, Patrícia Campos. Empresários bancam campanha contra o PT pelo WhatsApp. *Folha de São Paulo*, 18 out. 2018. Disponível em: <<https://www1.folha.uol.com.br/poder/2018/10/empresarios-bancam-campanha-contra-o-pt-pelo-whatsapp.shtml>> Acesso em 10 jul. 2020.

²² GOMES, Helton Simões; PADRÃO, Márcio. Seu contato vale R\$ 0,24: como funcionam as "máquinas de spam" no WhatsApp. *Tilt*, 20 out. 2018. Disponível em: <<https://www.uol.com.br/tilt/noticias/redacao/2018/10/20/como-funcionam-as-maquinas-de-distribuicao-em-massa-no-whatsapp.htm>> Acesso em 10 jul. 2020.

²³ GOMES, Helton Simões; PADRÃO, Márcio. Seu contato vale R\$ 0,24: como funcionam as "máquinas de spam" no WhatsApp. *Tilt*, 20 out. 2018. Disponível em: <<https://www.uol.com.br/tilt/noticias/redacao/2018/10/20/como-funcionam-as-maquinas-de-distribuicao-em-massa-no-whatsapp.htm>> Acesso em 10 jul. 2020.

²⁴ FERRARI, Pollyana; FILHO, Alberto Freitas. O Mundo está ao Contrário e Ninguém Reparou. In: TOURAL, Carlos; CORONEL, Gabriela; FERRARI, Pollyana (orgs.) *Big Data e Fake News na sociedade do (des)conhecimento*, 2. ed. Aveiro: Ria Editorial, 2020, p. 170.

James Susskind recorda que se alguém “quer que as pessoas odeiem seus inimigos, uma estratégia é disfarçar-se deles e dizer coisas repulsivas”²⁵, o que não é propriamente uma invenção moderna. Basta pensar nos “Protocolos de Sião”, que há séculos circulam como um meio de despertar o ódio contra os judeus. No entanto, Susskind chama a atenção para o fato de que os principais criadores de mentiras hoje são robôs que, graças à Inteligência Artificial, aprenderam a imitar o discurso humano. Estima-se, por exemplo, que só no ano de 2017, 48 milhões de usuários (9 a 15 por cento das contas) do Twitter eram *bots*.²⁶

Acerca dos riscos para a democracia, indaga Susskind:

A democracia deliberativa pode sobreviver em um sistema onde a própria deliberação não é mais um privilégio dos seres humanos? É possível que vozes humanas possam ser excluídas da esfera pública por *bots* que pouco se importam com nossas normas de conversação. No futuro, (...) elas podem parecer e soar como humanos, dotadas de rostos e vozes e extraordinários dons retóricos. Como podemos nós, com nossos cérebros fracos e conhecimento limitado, participar de forma significativa nas deliberações se nossas opiniões são instantaneamente despedaçadas por exércitos de *bots* armados com um milhão de réplicas espertinhas? Os defensores dos *bots* podem colocar de outra forma: por que perder tempo deliberando quando *bots* cada vez mais sofisticados podem debater as questões de maneira mais rápida e eficaz em nosso nome?²⁷

A discussão acerca da legalidade desse tipo de robô vem sendo amplamente enfrentada no Brasil com o Projeto de Lei (PL) das Fake News, definindo conta inautêntica como “criada ou usada com o propósito de assumir ou simular identidade de terceiros para enganar o público, ressaltados o direito ao uso de nome social e à pseudonímia nos termos desta Lei, bem como o explícito ânimo humorístico ou de paródia”, de acordo com art. 5º, II. Dessa forma, com o intuito de promover a liberdade de expressão, o acesso à informação e fluxo de ideias na internet, os provedores devem vedar a adoção de contas inautênticas (art. 6º, I, PL 2630/2020), podendo requerer a confirmação da identidade inclusive com envio de documento oficial de identidade (art. 7º, caput, PL 2630/2020)²⁸. Dessa forma, como pontua Carlos Affonso Souza, “[o] núcleo do conceito de comportamento inautêntico coordenado – que é proibido em muitas redes sociais – não está na coordenação, mas sim na autenticidade”²⁹, ou seja, no fato de um robô/bot com propósitos escusos de seus detentores estar se passando por uma pessoa.

²⁵ SUSSKIND, Jamie. *Future Politics: living together in a world transformed by tech*. Oxford: Oxford University Press, 2018, p. 232.

²⁶ Nesse sentido: “In the 2016 US presidential election, pro-Trump bots using hashtags like #LockHerUp flooded social media, outgunning the Clinton campaign’s own bots by 5:1 and spreading a whopping dose of fake news. It’s estimated that around one-third of all traffic on Twitter in the buildup to the EU Brexit referendum came from bots. Almost all were for the Leave side.” (SUSSKIND, Jamie. *Future Politics: living together in a world transformed by tech*. Oxford: Oxford University Press, 2018, p. 233).

²⁷ Tradução de: “Can Deliberative Democracy survive in a system where deliberation itself is no longer the preserve of human beings? It’s possible that human voices could be crowded out of the public sphere altogether by bots that care little for our conversational norms. In the future, (...) they could look and sound like humans, endowed with faces and voices and extraordinary rhetorical gifts. How can we, with our feeble brains and limited knowledge, participate meaningfully in deliberations if our views are instantaneously ripped to shreds by armies of bots armed with a million smart-ass retorts? Advocates of bots might put it differently: why spend time deliberating when increasingly sophisticated bots can debate the issues faster and more effectively on our behalf?” (SUSSKIND, Jamie. *Future Politics: living together in a world transformed by tech*. Oxford: Oxford University Press, 2018, p. 233).

²⁸ BRASIL. Senado Federal. *Projeto de Lei n.º 2630, de 03 de setembro de 2020*, Autor Alessandro Vieira, Institui a Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet. Brasília, Distrito Federal.

²⁹ AFFONSO, Carlos. O que difere a ação de perfis bolsonaristas no Facebook e de fãs do K-pop? *Tecfront*, 10 jul. 2020. Disponível em: <<https://tecfront.blogosfera.uol.com.br/2020/07/10/o-que-difere-a-acao-de-perfis-bolsonaristas-no-facebook-e-de-fas-do-k-pop/>> Acesso em 10 jul. 2020.

A Inteligência Artificial a Serviço da Desinformação: como as *Deepfakes* e as Redes Automatizadas Abalam a Liberdade de Ideias no Debate Público e a Democracia Constitucional e Deliberativa

Ao lado dessas redes automatizadas e das contas inautênticas, coloca-se o debate acerca das chamadas *deepfakes*³⁰, que se inserem num contexto mais amplo de reconstrução digital da imagem da pessoa, para criar vídeo-montagens com rostos quase idênticos aos reais, a partir de fotos e vídeos dos alvos. Exemplos não faltam no cinema, como é o caso da técnica usada “pelo diretor James Cameron no clássico *Avatar* para colocar o rosto dos atores Sam Worthington e Zoe Saldana nos gigantes azuis criados por computador.”³¹

Por definição, as *deepfakes* se referem a uma técnica baseada no *deep learning* (um aprofundamento do aprendizado de máquina característico da Inteligência Artificial)^{32 33}, capaz de criar vídeos falsos, a partir da troca do rosto de uma pessoa pelo de outra pessoa. A paternidade do termo é atribuída a um usuário do *site* Reddit com nome idêntico, que em 2017 alegou ter desenvolvido um algoritmo de *machine learning* que o ajudaria a colocar o rosto de celebridades em vídeos pornográficos. No entanto, como se pode ver, o uso, inicialmente atrelado à pornografia, seguiu seu rumo também pelos caminhos da desinformação e do humor.³⁴ Tãmanha é a perfeição alcançada pela tecnologia que seria necessária muita atenção para identificar se tratar de uma fraude, o que é especialmente perigoso nos tempos atuais, marcado pela “Economia da Atenção”.

Em princípio, as *deepfakes* são apenas de imagens e a voz da pessoa retratada normalmente é inserida por algum imitador. No entanto, já se está caminhando para o estágio do desenvolvimento tecnológico em que se torna possível manipular, também, a voz das pessoas.

Tãmanha é a gravidade das *deepfakes*, que a sua utilização para fins de pornografia de vingança (*revenge porn*) chegou a ser criminalizada por mais de uma dúzia de estados nos Estados Unidos da América.³⁵ O estado de Maryland foi além: debate uma lei que busca combater a influência política das *deepfakes*. A lei proibiria os indivíduos de “influenciar dolosa ou conscientemente ou de tentar influenciar a decisão de um eleitor de ir às urnas ou de votar em

³⁰ CHESNEY, Bobby; CITRON, Danielle. *Deep Fakes: A Looming Challenge for Privacy, Democracy and National Security* (Draft), pp. 31-41. Disponível em: <<https://ssrn.com/abstract=3213954>> Acesso em 10 jul. 2020.

³¹ SILVA, Rafael Rodrigues da. *Deepfakes no Brasil. Parte 2: a ameaça fantasma de nossa democracia. Canal Tech*, 24 out. 2019. Disponível em: <<https://canaltech.com.br/internet/deepfakes-no-brasil-parte-2-a-ameaca-fantasma-de-nossas-democracias-153453/>> Acesso em 10 jul. 2020.

³² Acerca dessa forma empregar a tecnologia: “Deep learning (DL) is the use of multilevel neural networks to find patterns in huge bodies of data (e.g. millions of images, or speech-sounds). The system isn’t told what patterns to look for, but finds them for itself.” (BODEN, Maggie. On deep learning, artificial neural networks, artificial life, and good-old fashioned AI. *Oxford University Press's Blog*. Disponível em: <<https://blog.oup.com/2016/06/artificial-neural-networks-ai/>> Acesso em 04 out. 2019). Mais sobre em: MEDON, Filipe. *Inteligência Artificial e Responsabilidade Civil: autonomia, riscos e solidariedade*. Salvador: JusPodivm, 2020, p. 56 e ss.

³³ Sobre essa técnica: “A técnica mais usada para esse tipo de vídeo é a chamada ‘troca de cabeças’, que consiste do uso de uma ‘pessoa-origem’ (a pessoa que você quer inserir no vídeo) e de uma ‘pessoa-destino’, cuja imagem será substituída pela da ‘pessoa-origem’. Assim, com o uso de softwares específicos que utilizam algoritmos de inteligência artificial (IA), é possível transferir o rosto da ‘pessoa-origem’ para o corpo da ‘pessoa-destino’ de forma que pareça que a ‘pessoa-origem’ realmente faz parte do vídeo, com uma dose assustadora de realismo.” (SILVA, Rafael Rodrigues da. *Deepfakes no Brasil. Parte 2: a ameaça fantasma de nossa democracia. Canal Tech*, 24 out. 2019. Disponível em: <<https://canaltech.com.br/internet/deepfakes-no-brasil-parte-2-a-ameaca-fantasma-de-nossas-democracias-153453/>> Acesso em 10 jul. 2020).

³⁴ TOLOSANA, Ruben; VERA-RODRIGUEZ, Ruben; FIERREZ, Julian; MORALES, Aythami; ORTEGA-GARCIA, Javier. *Deep Fakes and Beyond: A Survey of Face Manipulation and Fake Detection. ArXiv:2001.00179*. Disponível em: <<https://arxiv.org/abs/2001.00179>> Acesso em 10 jul. 2020.

³⁵ RUIZ, David. Deepfakes laws and proposals flood US. *Malwarebytes Labs*, 23 jan. 2020. Disponível em: <<https://blog.malwarebytes.com/artificial-intelligence/2020/01/deepfakes-laws-and-proposals-flood-us/>> Acesso em 10 jul. 2020.

um candidato em particular, publicando, distribuindo ou disseminando uma deepfake on-line dentro de 90 dias antes de uma eleição.”³⁶

Para fins de desinformação, o já narrado caso do governador de São Paulo João Dória serve para ilustrar a potencialidade lesiva de uma manipulação digital da imagem tão autêntica às vésperas de uma eleição. Até que a verdade seja apurada, a semente promíscua do mal já foi lançada e muitos eleitores, sobretudo aqueles mais indecisos, já podem ter mudado o voto em desfavor da vítima da imagem digitalmente reconstruída.

Um dos casos mais famosos da utilização de deepfakes para finalidades políticas foi o da jornalista Rana Ayyub na Índia:

Após a divulgação do caso do estupro de Kathua — uma menina de 8 anos que vivia na aldeia foi raptada, estuprada durante quase uma semana por quatro homens adultos, e então assassinada. Durante a apuração do caso, a liderança do partido BJP (partido de extrema direita da Índia e que defende ideais de supremacia hindu) partiu em defesa dos homens acusados do crime, alegando que aquilo era uma perseguição apenas por serem hindus — e ignorando que uma perícia conseguisse provas. Assim, Rana Ayyub foi uma das únicas figuras da imprensa que se colocou contra o partido (que hoje possui a maior parte das cadeiras do Congresso indiano e o apoio do primeiro-ministro Narendra Modi) e, em uma tentativa de silenciá-la, o BJP teria criado um vídeo pornô falso da jornalista utilizando técnicas de deepfake. Ainda que não existem provas contundentes de que o deepfake pornográfico foi uma criação do BJP, existem vários indícios de que isso seja verdade. Por exemplo, quem alertou Rana sobre a existência do vídeo foi um próprio membro do partido que se sentiu incomodado com a tática. Quando esse membro alertou a jornalista sobre a existência do vídeo, ele ainda estava sendo compartilhado apenas nos grupos de WhatsApp dos membros do BJP, mas horas depois, a própria página oficial do Facebook do líder do BJP compartilhou o vídeo na rede social, o que deixou claro que, mesmo que o conteúdo não tenha sido criado dentro do partido, ele não só apoiava a criação como ainda o compartilhamento deste conteúdo como forma de tirar a credibilidade e silenciar um adversário político.³⁷

Como se pode observar, o emprego de técnicas de Inteligência Artificial para desinformar, seja por meio de redes automatizadas, contas inautênticas e *deepfakes* sinaliza para uma importante conclusão: a necessidade de se controlar não o conteúdo do discurso que promove a desinformação, mas a sua autenticidade e a sua atividade. Dito em outras palavras: é muito mais fácil e eficaz rastrear e dismantelar redes de desinformação pela identificação de atividade anormal, claramente derivada de tecnologia robótica, do que pelo conteúdo que ela venha a produzir. Isso ressalta a importância de se utilizar a própria tecnologia para rastrear os usos anormais da própria tecnologia, conforme será visto no item abaixo.

3. Impactos na Democracia Constitucional e nas Liberdades Comunicacionais

A democracia e a soberania popular são elementos centrais do constitucionalismo contemporâneo. Assim, a “constituição não só deve ter origem democrática, como deve

³⁶ No original: “On January 16, Maryland introduced a bill targeting political influence deepfakes. The bill, which has a scheduled hearing in early February, prohibits individuals from “willfully or knowingly influencing or attempting to influence a voter’s decision to go to the polls or to cause a vote for a particular candidate by publishing, distributing, or disseminating a deepfake online within 90 days of an election.” (RUIZ, David. Deepfakes laws and proposals flood US. In: *Malwarebytes Labs*, 23 jan. 2020. Disponível em: <<https://blog.malwarebytes.com/artificial-intelligence/2020/01/deepfakes-laws-and-proposals-flood-us/>> Acesso em 10 jul. 2020.) Para acompanhar a discussão legislativa: <<https://www.billtrack50.com/BillDetail/1174065>> Acesso em 10 jul. 2020.

³⁷ SILVA, Rafael Rodrigues da. Deepfakes no Brasil. Parte 2: a ameaça fantasma de nossa democracia. In: *Canal Tech*, 24 out. 2019. Disponível em: <<https://canaltech.com.br/internet/deepfakes-no-brasil-parte-2-a-ameaca-fantasma-de-nossas-democracias-153453/>> Acesso em 10 jul. 2020.

A Inteligência Artificial a Serviço da Desinformação: como as *Deepfakes* e as Redes Automatizadas Abalam a Liberdade de Ideias no Debate Público e a Democracia Constitucional e Deliberativa

organizar um Estado que assegure a soberania popular. (...). O princípio democrático serve como legitimador da constituição no sentido da congruência entre a soberania popular e o Estado”³⁸. O sentimento popular de dissociação entre, de um lado, agentes e órgãos públicos, e, de outro lado, cidadãos é grande nas democracias constitucionais. Apesar das vantagens da representação política apontada pelos Federalistas norte-americanos, a qual seria relevante para permitir uma deliberação mais equilibrada, a pequena responsividade popular dos agentes políticos e a ausência de representação da pluralidade social colocam em xeque a real dimensão democrática no constitucionalismo contemporâneo. Dessa forma, “o sistema representativo nasceu prometendo-nos algo que hoje não está em condições de cumprir”³⁹.

Nesse contexto, nos âmbitos da filosofia política e do direito constitucional, retomam-se com ênfase as reflexões sobre a deliberação popular de temas sociais relevantes e centrais, repensando assim as formas pelas quais as estruturas do Estado Direito e especialmente os Poderes estatais devem cooperar com a soberania popular. A democracia deliberativa e constitucional está comprometida com os direitos fundamentais e especialmente com os direitos fundamentais à liberdade e à igualdade, os quais efetivam a soberania popular nos marcos no constitucionalismo. Nesse sentido, Cláudio Pereira de Souza Neto afirma que:

As liberdades protegidas pela democracia deliberativa não são apenas aquelas que exercem uma função imediata no processo democrático. A autonomia privada, em sua dimensão não imediatamente política, também está incluída, pois as pessoas só se dispõem a cooperar quando seus projetos de vida razoáveis são tratados como dignos de igual respeito. (...)

A democracia deliberativa exige não só a igualdade de “possibilidades” para participar da vida pública, mas também igualdade de “capacidades” para fazê-lo efetivamente (...)⁴⁰

Por sua vez, a utilização das novas tecnologias, de redes sociais, dos aplicativos e da internet altera substancialmente a deliberação pública e a pauta republicana nas democracias contemporâneas. Por exemplo, há o fenômeno das “bolhas” nas interações sociais pelas redes sociais e atualmente por meio de aplicativos de mensagens, reduzindo substancialmente a pluralidade e a diversidade nas interações comunicativas⁴¹. As novas tecnologias permitem que o

³⁸ BERCOVICI, Gilberto. *Soberania e Constituição: para uma Crítica do Constitucionalismo*. São Paulo: Quartier Latin, 2008, p. 19.

³⁹ Tradução livre de: “el sistema representativo nació prometiéndonos algo que hoy no se encuentra en condiciones de cumplir” (GARGARELLA, Roberto. *Crítica de La Constitución: sus Zonas Oscuras*. Buenos Aires: Claves para Todos, 2004, p. 19).

⁴⁰ NETO SOUZA, Cláudio Pereira de. *Teoria Constitucional e Democracia Deliberativa: um Estudo sobre Papel do Direito na Garantia das Condições para a Cooperação na Deliberação Democrática*. Rio de Janeiro: Renovar, 2006, p. 296-297.

⁴¹ As “bolhas digitais” podem ser validamente definidas como o *locus* virtual de confirmação de vieses cognitivos, onde não há espaço para a construção dialética de ideias e a todo tempo une e reafirma concepções preexistentes em grupos que conectam pessoas em torno dos mesmos preconceitos, formando verdadeiras “ilhas” de pensamentos isolados e imunes ao debate. Nesse sentido, Vania Baldi explica que: “De facto, numerosas investigações (Bikhchandani, Hirshleifer, Welch, 1998; Hirshleifer, 1999) refletem sobre os vieses de confirmação que acompanham as contraposições especulares entre “progressistas” e “conservadores” sobre, por exemplo, as temáticas ambientais, médicas e energéticas em contextos pré-digitais.⁴¹ Nestes casos podemos considerar que não são as tecnologias digitais que criam as condições de enviesamento dos factos, uma vez que estes são o resultado de preconceitos e influências sociais alheias ao mundo digital. As tecnologias digitais proporcionam a radicalização de posições contrapostas já existentes, assim como o desincentivo à auscultação recíproca e a indiferença às respetivas razões expressas e aos alegados conhecimentos. Em poucas palavras, o mundo interconectado revela-se (num sentido sociotécnico) inesperadamente alérgico ao pluralismo e produtor de uma série alargada de *networked individualism* (Welman et al, 2003)” (BALDI, Vania. *A construção viral da realidade: ciberpopulismo e polarização dos públicos em rede*. Observatorio (OBS*) *SpecialIssue*, (2018), 004-020 Disponível em: <http://obs.obercom.pt/index.php/obs/article/view/1420/pdf>. Acesso 03/08/2019).

indivíduo escolha os fóruns que deseja engajar-se e as pessoas com as quais busca dialogar, fazendo com que os debates e as comunicações se encontrem restritas às áreas de interesse individual. A lógica adotada passa a ser da comunicação-consumo, deixando em segundo plano as preocupações e práticas republicanas.⁴²

A internet, os aplicativos e as redes sociais potencializaram a radicalização dos discursos e diálogos entre pessoas que pensam iguais, não produzindo um fórum público republicano na internet ou por meio de aplicativo, e sim uma enorme fragmentação social com grande dificuldade para permitir que diversas visões e tendências participem de debates sociais. Em síntese, ocorre qualitativamente o empobrecimento das reflexões, análises e discussões sociais⁴³. Ainda, o modelo de negócio das redes sociais impulsiona a segmentação dos grupos sociais e a degradação do espaço público, que são “potencializadas extraordinariamente pelas redes sociais, (...). O advento da instabilidade política e dos conflitos sociais virtuais através das redes incrementa os recursos publicitários percebidos”⁴⁴.

Esses problemas sobre a deliberação nas democracias contemporâneas são ampliados com as ações coordenadas inautênticas e as *deepfakes*, dificultando ainda mais a implementação de um modelo robusto de democracia deliberativa centrado na liberdade e na igualdade materiais. No mundo digital, a liberdade é equivalente ao poder de escolha do consumidor sobre os produtos comunicativos que pretende adquirir, passando o indivíduo, por sua vez, a receber uma enorme quantidade de publicidade em razão dos seus cliques e do modelo de negócio adotado pelas redes sociais e pelas plataformas eletrônicas. Essa perspectiva privatista da liberdade em muito difere das propostas da democracia participativa que pretendem implementar mecanismos e instrumentos de efetiva participação popular, a partir de informações e conhecimentos suficientes e verossímeis⁴⁵.

As *deepfakes*, as quais são utilizadas de forma deliberada para enganar os demais indivíduos com conteúdos falsos com aparência de veracidade, e as ações coordenadas inautênticas são claros elementos desinformação e obviamente não estão protegidos pelo âmbito de proteção da liberdade de expressão, já que não são corolários do livre desenvolvimento da personalidade dos seres humanos, assim como não auxiliam, e sim corroem a liberdade de ideias na esfera pública. Sobre o âmbito de proteção da liberdade de expressão:

abarca um conjunto diferenciado de situações, cobrindo, em princípio, uma série de liberdades (faculdades) de conteúdo espiritual, incluindo expressões não verbais, como é o caso da expressão musical, da comunicação pelas artes plásticas e dentre outras. A liberdade de expressão consiste, mais precisamente, na liberdade de exprimir opiniões, ou seja, juízos de valor a respeito de fatos, ideias, opiniões de terceiros, etc⁴⁶.

Por excelência as *deepfakes* criam expressões verbais e não verbais falsas que aparentemente foram emitidas por certos sujeitos com objetivo de degradar essas pessoas e seus movimentos, beneficiando assim opositores e grupos adversários de forma dissimulada e rasteira. A liberdade de expressão consiste no poder de a pessoa humana “adotar a atitude intelectual de

⁴² Sobre o tema, cf. SUNSTEIN, Cass R. *Republic.com 2.0*. Princeton: Princeton University Press, 2007, p. 46-96 e CALLEJÓN, Francisco Balaguer. As duas grandes crises do constitucionalismo diante da globalização no século XXI. *Espaço Jurídico Journal of Law*, 19 (3), 681-702, 2019.

⁴³ SUNSTEIN, Cass R. *Republic.com 2.0*. Princeton: Princeton University Press, 2007, p. 43-45.

⁴⁴ CALLEJÓN, Francisco Balaguer. As duas grandes crises do constitucionalismo diante da globalização no século XXI. *Espaço Jurídico Journal of Law*, 19 (3), 684, 2019.

⁴⁵ Sobre a democracia participativa, cf. COSTA, Pietro. *Poucos, Muitos, Todos: Lições de História da Democracia*. Curitiba: Editora UFPR, 2012, p. 299-307.

⁴⁶ SARLET, Ingo Wolfgang; ROBL FILHO, Ilton Norberto. Estado Democrático de Direito e os limites da liberdade de expressão na constituição federal de 1988, com destaque para o problema da sua colisão com outros direitos fundamentais, em especial, com os direitos de personalidade. *Constituição, Economia e Desenvolvimento: Revista da Academia Brasileira de Direito Constitucional*. Curitiba, 2016, vol. 8, n. 14, Jan.-Jun. p. 119.

A Inteligência Artificial a Serviço da Desinformação: como as *Deepfakes* e as Redes Automatizadas Abalam a Liberdade de Ideias no Debate Público e a Democracia Constitucional e Deliberativa

sua escolha: quer um pensamento íntimo, quer seja a tomada de posição pública; liberdade de pensar e dizer o que se crê verdadeiro”⁴⁷. De outro lado, as *deepfakes* difundem uma atitude ou um pensamento que não foi adotado ou escolhido pelo suposto emissor, tentando fazer crer aos receptores da comunicação que a atitude, o pensamento e a expressão eram verdadeiros. Esse uso tecnológico escuso não está previsto na liberdade de expressão, e sim afronta o âmbito de proteção desse direito fundamental.

As contas inautênticas, de outro lado, potencializam artificialmente os emissores de certas opiniões e reflexões no mercado de ideias⁴⁸, majorando o poder de convencimento e de persuasão sobre os demais participantes da comunicação. Nesse caso, a desigualdade informacional é ampliada, inflando a quantidade de indivíduos que sustentam certas posições e causando o efeito manada para diversos receptores dessa comunicação, os quais muitas vezes passam a acreditar em razão do exponencial número de emissores de certa opinião ou reflexão. Ainda esses movimentos coordenados inautênticos, com frequência, disseminam informações falsas, incorretas e errôneas, que turbam e conduzem a erros boa parte dos receptores desses conteúdos.

Dessa forma, tanto as *deepfakes* como as contas inautênticas ampliam a degradação sobre uma esfera pública igualitária, livre e plural, além de atentar contra o direito à informação especialmente na sua dimensão difusa, porque existe necessidade de controle “da veracidade das informações veiculadas, tutelando o direito difuso da sociedade de recebê-las de maneira isenta e não fraudulenta. Não se trata, propriamente, da busca de uma verdade absoluta, mas, tão somente, o controle da qualidade de ser veraz”⁴⁹. Obviamente a análise sobre a veracidade da informação é tema complexo principalmente pelos possíveis impactos sobre as liberdades comunicacionais e constitucionais, nos termos do art. 5º, IV, VI e IX, Constituição Federal brasileira de 1988.

As liberdades de expressão e comunicacional possuem preferência condicionada em relação à tutela do direito difuso de receber informação verossímil, pois as referidas liberdades são imprescindíveis para a construção social democrática e para qualquer esfera pública pluralista, assim como representam requisitos essenciais do livre desenvolvimento da personalidade, de acordo com art. 1º, caput, III e V, CF/88⁵⁰. Dessa maneira, eventual análise sobre o conteúdo deve ser excepcional nas redes sociais e nas interações digitais, focando o controle “no reconhecimento de padrões de comportamento do que no conteúdo em si. Isso reduz o risco de estarmos automatizando a censura”⁵¹. As plataformas de informação e os aplicativos precisam tomar medidas de controle especialmente de contas inautênticas, mas também das *deepfakes*, já

⁴⁷ SILVA, José Afonso da. *Curso de Direito Constitucional Positivo*. 28. ed., rev. e atualizada. São Paulo: Malheiros, 2007, p. 241.

⁴⁸ Sobre o livre mercado de ideias, o qual se trata originalmente de construção da jurisprudência norte-americana e posteriormente disseminada, cf. MACHADO, Jónatas E.M. *Liberdade de Expressão: Dimensões Constitucionais da Esfera Pública no Sistema Social*. Coimbra: Coimbra Editora, 2002, p. 237-247. Para uma análise crítica dos limites do livre mercado de ideias, cf. SUNSTEIN, Cass R. *Democracy and the Problem of Free Speech*. New York: The Free Press, 1995.

⁴⁹ CARVALHO, Luis Gustavo Grandinetti Castanho de. O Ministério Público e a Mídia. *Constituição, Economia e Desenvolvimento*: Revista da Academia Brasileira de Direito Constitucional. Curitiba, vol. 7, n. 12, Jan.-Jun. p. 39, 2015.

⁵⁰ Sobre os fundamentos da liberdade de expressão com ênfase ao aspecto democrático, cf. ROBL FILHO, Ilton Norberto. Fundamentos Filosóficos e Políticos das Liberdades de Imprensa e de Expressão no Estado Democrático de Direito: Algumas Reflexões a Partir do Constitucionalismo Norte-Americano e do Supremo Tribunal Federal Brasileiro na ADPF nº. 130. In: ARAÚJO, Marcelo Labanca Corrêa de; MAIA, Luciano Mariz; SILVA, Lucas Gonçalves da.. (Org.). *Direitos Fundamentais e Democracia I*; XXIII Encontro Nacional do CONPEDI. 1ed. Florianópolis: CONPEDI, 2014, v. 1, p. 337-353.

⁵¹ AFFONSO, Carlos. O que difere a ação de perfis bolsonaristas no Facebook e de fãs do K-pop? *Tecfront*, 10 jul. 2020. Disponível em: <<https://tecfront.blogosfera.uol.com.br/2020/07/10/o-que-difere-a-acao-de-perfis-bolsonaristas-no-facebook-e-de-fas-do-k-pop/>> Acesso em 10 jul. 2020.

que são fortes instrumentos de desinformação com enorme impacto nas interações comunicativas.

As atuações inautênticas nos espaços virtuais e os diversos mecanismos digitais que promovem a desinformação auxiliam na promoção do populismo digital⁵² e na concretização de contemporâneas democraturas. As informações são transmitidas especialmente por meio de acessos gratuitos dos receptores/leitores da internet e de aplicativos, possibilitando, de outro lado, que qualquer pessoa envie com rapidez conteúdos, reflexões e opiniões. Com frequência, desinformação é espalhada e contas inautênticas divulgam comunicações que, em geral, beneficiam grupos políticos e líderes populistas, os quais utilizam com habilidade conteúdos falsos para obter apoio. Esses movimentos populistas trabalham claramente com a lógica amigo/inimigo, a partir da dicotomia líderes populares e “povo” contra supostas elites política, econômica, intelectual e política.

O populismo digital desconsidera a pluralidade social. Ao contrário da visão autoritária do populismo, inexistente um conceito homogêneo de povo, o qual comunga de forma unívoca opiniões, visões e propostas sociais, pois a diversidade é a marca característica da democracia constitucional, logo igualitária, livre, inclusiva e deliberativa⁵³. O líder e o movimento populista empregam a desinformação e a utilização de contas inautênticas com frequência para manter ou ampliar o apoio de parte da população sobre a suposta pauta nacional do povo de bem contra as elites corrompidas e autointeressadas.

O populismo com a utilização distorcida dos meios digitais, quando se encontra no poder, é terreno fértil para democraturas, as quais são regimes formalmente democráticos, mas que se apropriam de instrumentos constitucionais de forma deturpada para manutenção do poder, estabelecendo como regra medidas autoritárias. Esse fenômeno é bem estudado pela literatura especializada⁵⁴, devendo a teoria constitucional estar atenta e vigilante para combater as práticas autoritárias.

4. Conclusão

Em texto clássico, Rudolf Smend afirma que a verdadeira substância da Constituição “não é uma existência imóvel (...), mas simplesmente vida que flui, se renova constantemente”⁵⁵. Os dilemas sobre a liberdade de expressão e o direito à informação alteram-se substancialmente em contexto de utilização massiva da internet, dos aplicativos, das redes sociais e das plataformas eletrônicas, além de que a esfera pública fragmenta-se com redução da pluralidade, porque não se observa com frequência pessoas e grupos diversos debatendo de forma minimamente construtiva no mesmo espaço digital.

O fenômeno das bolhas amplia a valorização do indivíduo-consumidor, permitindo escolher as discussões e os grupos que vai participar a partir de interesses particulares e pessoais na internet, aplicativos e redes sociais. Essa forma de debate como regra produz a radicalização

⁵² Sobre essa questão, especialmente cf. LAGO, Alessandro dal. *Populismo Digitale: La Crisi, la Rete e la Nuova Destra*. Milano: Raffaello Cortina Editore, 2017, p. 9-29.

⁵³ Para aprofundamento, cf. MARRAFON, Marco Aurélio. PANSIERI, Flávio. ROBL FILHO, Ilton Norberto. Constitucionalismo como Salvaguarda do Estado de Direito: Crítica ao (Ciber)populismo Autoritário e a Necessária Reengenharia Constitucional. *Revista do Instituto de Hermenêutica Jurídica*, Belo Horizonte, n. 28, jul/dez 2020.

⁵⁴ Sobre o tema, cf. SCHEPPELE, Kim Lane. The Opportunism of Populists and the Defense of Constitutional Liberalism. *German Law Journal*, 20(3), 314-331, 2019.

⁵⁵ Tradução livre de: “no es una existencia inmóvil (...), sino simplemente vida que fluye, que se renueva constantemente” (SMEND, Rudolf. *Ensayos sobre la Libertad de Expresión, de Ciencia y de Cátedra como Derecho Fundamental y sobre el Tribunal Constitucional Federal Alemán*. México D. F.: Universidad Nacional Autónoma de México, 2005, p. 6.

A Inteligência Artificial a Serviço da Desinformação: como as *Deepfakes* e as Redes Automatizadas Abalam a Liberdade de Ideias no Debate Público e a Democracia Constitucional e Deliberativa

dos discursos e a pequena possibilidade de construção de acordos e de soluções coletivas dialogadas, pois os pensamentos diversos não se encontram e, quando há interação de membros das bolhas entre si, as comunicações em geral são xingamentos ou ataques imputando estereótipos negativos aos diversos.

As redes automatizadas que utilizam robôs e contas inautênticas em aplicativos e na internet ampliam os problemas nas comunicações eletrônicas e digitais, já que são com frequência instrumentos de produção e de disseminação de informações falsas. Ainda quando não criam e publicizam conteúdos falsos, as contas inautênticas promovem a distorção do livre mercado de ideias, em razão da majoração exponencial de emissores sobre determinadas visões, distorcendo no mínimo as forças no processo de discussão e deliberação. Por sua vez, as *deepfakes* com exceção da sua utilização para o humor têm por função promover a desinformação, atacando de maneira maliciosa e injusta pessoas e grupos. Tanto as *deepfakes* como as contas inautênticas não se encontram albergadas pelos âmbitos de proteção da liberdade de expressão e do direito à informação. De outro lado, o combate a essas distorções comunicativas deve ser feito com cautela, privilegiando como regra a verificação de padrões inautênticos e de aspectos técnicos e tecnológicos.

Por fim, líderes e grupos populistas trabalham com desenvoltura e frequência por meio dos instrumentos de comunicação das redes sociais, aplicativos e internet. Dessa forma e pela possibilidade desses grupos, quando no poder, produzem transições espúrias de regimes democráticos para regimes formalmente democráticos, mas materialmente autoritários, o direito constitucional precisa estar atento e atuante contra os ataques ao Estado Democrático de Direito.

6. Referências

AFFONSO, Carlos. Tecnologia abre novo capítulo na manipulação de vídeos; imagina na eleição. *Tecfront*, 28 mai. 2019. Disponível em: <<https://tecfront.blogosfera.uol.com.br/2019/05/28/tecnologia-abre-novo-capitulo-na-manipulacao-de-videos-imagina-na-eleicao/>> Acesso em 10 abr. 2020.

_____. O que difere a ação de perfis bolsonaristas no Facebook e de fãs do K-pop? *Tecfront*, 10 jul. 2020. Disponível em: <<https://tecfront.blogosfera.uol.com.br/2020/07/10/o-que-difere-a-acao-de-perfis-bolsonaristas-no-facebook-e-de-fas-do-k-pop/>> Acesso em 10 jul. 2020.

BALDI, Vania. *A construção viral da realidade: ciberpopulismo e polarização dos públicos em rede*. Observatorio (OBS*) *SpecialIssue*, (2018), 004-020 Disponível em: <<http://obs.obercom.pt/index.php/obs/article/view/1420/pdf> . > Acesso 03/08/2019.

BALKIN, Jack. The path of robotics law. *California Law Review Circuit*, Berkeley, v. 06, p. 45-60, jun. 2015.

BENKLER, Yochai; FARIS, Robert; ROBERTS, Hal. *Network propaganda: manipulation, disinformation, and radicalization in American politics*. New York: Oxford University Press, 2018.

BODEN, Maggie. On deep learning, artificial neural networks, artificial life, and good-old fashioned AI. *Oxford University Press's Blog*. Disponível em: <<https://blog.oup.com/2016/06/artificial-neural-networks-ai/>> Acesso em 04 out. 2019.

BRASIL. Senado Federal. *Projeto de Lei n.º. 2630, de 03 de setembro de 2020*, Autor Alessandro Vieira, Institui a Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet. Brasília, Distrito Federal.

BERCOVICI, Gilberto. *Soberania e Constituição: para uma Crítica do Constitucionalismo*. São Paulo: Quartier Latin, 2008.

CALLEJÓN, Francisco Balaguer. As duas grandes crises do constitucionalismo diante da globalização no século XXI. *Espaço Jurídico Journal of Law*, 19 (3), p. 681-701, 2019.

CARVALHO, Luis Gustavo Grandinetti Castanho de. O Ministério Público e a Mídia. *Constituição, Economia e Desenvolvimento: Revista da Academia Brasileira de Direito Constitucional*. Curitiba, vol. 7, n. 12, Jan.-Jun. p. 33-50, 2015.

CHESNEY, Bobby; CITRON, Danielle. *Deep Fakes: A Looming Challenge for Privacy, Democracy and National Security (Draft)*, pp. 31-41. Disponível em: <<https://ssrn.com/abstract=3213954>> Acesso em 10 jul. 2020.

CHOKSHI, Niraj. That Wasn't Mark Twain: How a Misquotation Is Born. In: *The New York Times*, 26 abr. 2017. Disponível em: <<https://www.nytimes.com/2017/04/26/books/famous-misquotations.html>> Acesso em 08 jul. 2020.

COSTA, Pietro. *Poucos, Muitos, Todos: Lições de História da Democracia*. Curitiba: Editora UFPR, 2012.

FOLHA DE SÃO PAULO. Peritos constataram montagem em vídeo vazado, afirma Doria. 24 out. 2018. Disponível em: <<https://www1.folha.uol.com.br/poder/2018/10/peritos-constataram-montagem-em-video-vazado-afirma-doria.shtml>> Acesso em 09 jul. 2020.

FERRARI, Pollyana; FILHO, Alberto Freitas. O Mundo está ao Contrário e Ninguém Reparou. In: TOURAL, Carlos; CORONEL, Gabriela; FERRARI, Pollyana (orgs.) *Big Data e Fake News na sociedade do (des)conhecimento*, 2. ed. Aveiro: Ria Editorial, 2020, p. 153-173.

G1. *Facebook remove rede de contas falsas relacionada ao PSL e a gabinetes da família Bolsonaro*. 08 jul. 2020. Disponível em: <<https://g1.globo.com/economia/tecnologia/noticia/2020/07/08/facebook-remove-rede-de-contas-falsas-relacionada-ao-psl-e-a-gabinetes-da-familia-bolsonaro.ghtml>> Acesso em 08 jul. 2020.

GARGARELLA, Roberto. *Crítica de La Constitución: sus Zonas Oscuras*. Buenos Aires: Claves para Todos, 2004.

GOMES, Helton Simões; PADRÃO, Márcio. Seu contato vale R\$ 0,24: como funcionam as "máquinas de spam" no WhatsApp. *Tilt*, 20 out. 2018. Disponível em: <<https://www.uol.com.br/tilt/noticias/redacao/2018/10/20/como-funcionam-as-maquinas-de-distribuicao-em-massa-no-whatsapp.htm>> Acesso em 10 jul. 2020.

LAGO, Alessandro dal. *Populismo Digitale: la Crisi, la Rete e la Nuova Destra*. Milano: Raffaello Cortina Editore, 2017.

A Inteligência Artificial a Serviço da Desinformação: como as Deepfakes e as Redes Automatizadas Abalam a Liberdade de Ideias no Debate Público e a Democracia Constitucional e Deliberativa

MACHADO, Jónatas E.M. *Liberdade de Expressão: Dimensões Constitucionais da Esfera Pública no Sistema Social*. Coimbra: Coimbra Editora, 2002.

MARRAFON, Marco Aurélio. PANSIERI, Flávio. ROBL FILHO, Ilton Norberto. Constitucionalismo como Salvaguarda do Estado de Direito: Crítica ao (Ciber)populismo Autoritário e a Necessária Reengenharia Constitucional. *Revista do Instituto de Hermenêutica Jurídica*, Belo Horizonte, n. 28, jul/dez 2020.

MEDON, Filipe. *Inteligência Artificial e Responsabilidade Civil: autonomia, riscos e solidariedade*. Salvador: JusPodivm, 2020.

MELLO, Patrícia Campos. Empresários bancam campanha contra o PT pelo WhatsApp. *Folha de São Paulo*, 18 out. 2018. Disponível em: <<https://www1.folha.uol.com.br/poder/2018/10/empresarios-bancam-campanha-contr-o-pt-pelo-whatapp.shtml>> Acesso em 10 jul. 2020.

NETO SOUZA, Cláudio Pereira de. *Teoria Constitucional e Democracia Deliberativa: um Estudo sobre Papel do Direito na Garantia das Condições para a Cooperação na Deliberação Democrática*. Rio de Janeiro: Renovar, 2006.

RUIZ, David. Deepfakes laws and proposals flood US. *Malwarebytes Labs*, 23 jan. 2020. Disponível em: <<https://blog.malwarebytes.com/artificial-intelligence/2020/01/deepfakes-laws-and-proposals-flood-us/>> Acesso em 10 jul. 2020.

[ROBL FILHO, Ilton Norberto](#). Fundamentos Filosóficos e Políticos das Liberdades de Imprensa e de Expressão no Estado Democrático de Direito: Algumas Reflexões a Partir do Constitucionalismo Norte-Americano e do Supremo Tribunal Federal Brasileiro na ADPF nº. 130. In: ARAÚJO, Marcelo Labanca Corrêa de; MAIA, Luciano Mariz; SILVA, Lucas Gonçalves da.. (Org.). *Direitos Fundamentais e Democracia I*; XXIII Encontro Nacional do CONPEDI. Florianópolis: CONPEDI, 2014, v. 1, p. 337-353.

SARLET, Ingo Wolfgang; ROBL FILHO, Ilton Norberto. Estado Democrático de Direito e os limites da liberdade de expressão na constituição federal de 1988, com destaque para o problema da sua colisão com outros direitos fundamentais, em especial, com os direitos de personalidade. *Constituição, Economia e Desenvolvimento: Revista da Academia Brasileira de Direito Constitucional*. Curitiba, vol. 8, n. 14, Jan.-Jun., p. 112-142, 2016.

SCHEPPELE, Kim Lane. The Opportunism of Populists and the Defense of Constitutional Liberalism. *German Law Journal*, 20(3), 314-331, 2019.

SILVA, José Afonso da. *Curso de Direito Constitucional Positivo*. 28. ed., rev. e atualizada. São Paulo: Malheiros, 2007.

SILVA, Rafael Rodrigues da. Deepfakes no Brasil. Parte 2: a ameaça fantasma de nossa democracia. *Canal Tech*, 24 out. 2019. Disponível em: <<https://canaltech.com.br/internet/deepfakes-no-brasil-parte-2-a-ameaca-fantasma-de-nossas-democracias-153453/>> Acesso em 10 jul. 2020.

SMEND, Rudolf. *Ensayos sobre la Libertad de Expresión, de Ciencia y de Cátedra como Derecho Fundamental y sobre el Tribunal Constitucional Federal Alemán*. México D. F.: Universidad Nacional Autónoma de México, 2005.

SOUZA, Carlos Affonso Pereira de; TEFFÉ, Chiara Spadaccini de. Fake news: como garantir liberdades e conter notícias falsas na internet? In: MENEZES, Joyceane Bezerra de; TEPEDINO, Gustavo (Coord.), *Autonomia Privada, liberdade existencial e direitos fundamentais*. Belo Horizonte: Fórum, 2019, pp. 525-543.

SUNSTEIN, Cass R. *Democracy and the Problem of Free Speech*. New York: The Free Press, 1995.

_____. *Republic.com 2.0*. Princeton: Princeton University Press, 2007, p. 46-96.

SUSSKIND, Jamie. *Future Politics: living together in a world transformed by tech*. Oxford: Oxford University Press, 2018

TOLOSANA, Ruben; VERA-RODRIGUEZ, Ruben; FIERREZ, Julian; MORALES, Aythami; ORTEGA-GARCIA, Javier. *Deep Fakes and Beyond: A Survey of Face Manipulation and Fake Detection*. *ArXiv:2001.00179*. Disponível em: <<https://arxiv.org/abs/2001.00179>> Acesso em 10 jul. 2020.