

**“Só acredito vendo”:  
as *deepfakes* como mais uma forma de desinformar**

***“I’ll only believe it when I see it”:  
deepfakes as another form of disinformation***

Luiz Ricardo Goulart HÜTTNER<sup>1</sup>

### Resumo

O presente artigo tem como objetivo trazer a discussão de como as *deepfakes* são usadas para modificar falas e situações, a princípio, no Brasil, com o intuito de divertir, mas uma ferramenta que tem condições e potencial para manchar reputações e fazer com que a sociedade seja moldada para aquilo que os produtores pretendem, ou seja, na desinformação. Assim, buscamos fazer um breve relato sobre o que está sendo pesquisado no mundo sobre o tema e trazer um exemplo que é popular no Brasil, podendo ter um potencial grande para desinformar.

**Palavras-chave:** *Deepfake*. Desinformação. *Fake News*. Vídeo.

### Abstract

This article aims to bring up the discussion of how deepfakes are used to modify speeches and situations, at first, in Brazil, in order to have fun, but a tool that has conditions and potential to tarnish reputations and make society be shaped for what the producers want, that is, in disinformation. Thus, we seek to make a brief report on what is being researched in the world on the subject and to bring an example that is popular in Brazil, which may have great potential to disinform.

**Keywords:** Deepfake. Disinformation. Fake News. Video.

### Introdução

O presente artigo faz parte do conjunto de esforços, no Brasil, para delimitarmos as categorias de desinformação presentes na língua portuguesa. Os trabalhos acadêmicos que versam sobre *fake news* estão atualmente crescendo em nossa língua,

---

<sup>1</sup> Doutorando do Programa de Pós-Graduação em Comunicação da Universidade Federal do Rio Grande do Sul (UFRGS). Bolsista CAPES. E-mail: luizricardohuttner@gmail.com

mas muito do que é estudado em outras línguas é incorporado em nossas pesquisas, mesmo assim o termo é amplamente utilizado em diversas mídias (como televisão e rádio, por exemplo). Dessa forma, buscamos esclarecer um fenômeno que, ainda, não se tem pesquisas profundas, mas que se mostram com um potencial para a desinformação, são as *deepfakes*. Muito comum na sociedade falarmos que acreditamos naquilo que vemos e as imagens fortalecem essa visualização, mas será que tudo o que vemos é verdadeiro, real e factual?

As *Deepfakes*, em resumo, são vídeos (ou imagens) modificados que, se utilizados para a desinformar, buscam manchar reputações de pessoas ou de instituições, especialmente aquelas ligadas a política ou movimentos políticos. É importante recapitularmos o que vem sendo desenvolvido e pesquisado quanto a desinformação, chegando ao fenômeno que estamos começando a conhecer. Assim, temos como objetivo analisar como se dá as *deepfakes* no Brasil e fazer um referencial teórico sobre o tema que vem sendo tratado ao redor do mundo, além de trazermos exemplos de como isso tem se espalhado nas mídias sociais. Utilizaremos o exemplo de como o humor é usado para a modificação de imagens com o intuito (neste caso) de entreter, mas que facilmente pode ser utilizado para outros fins.

### **Modificações na comunicação: o potencial para desinformar**

O mundo passou, ao longo dos anos, por muitas modificações tecnológicas em todas as suas áreas, modificações que aconteceram e foram incorporadas ao cotidiano do ser humano. Junto com o acesso das pessoas a essas novas tecnologias (e pela história passamos pelo surgimento da fotografia, do cinema, dos jornais impressos e suas evoluções, por exemplo) fizeram com que a mídia (e todas as suas evoluções) tivesse presente cada vez mais na vida em sociedade. Já em 1999, por exemplo, Castells afirmava que

a Internet tem tido um índice de penetração mais veloz do que qualquer outro meio de comunicação na história: nos Estados Unidos, o rádio levou trinta anos para chegar aos sessenta milhões de pessoas; a TV alcançou esse nível de difusão em 15 anos; a Internet o fez em apenas três anos após a criação da teia mundial (CASTELLS, 1999, p. 439).

E isso se aplica muito a comunicação que temos e conhecemos hoje. Segundo Pavlik (2005), a mudança que a internet trouxe a comunicação são inúmeras, como: notícias onipresentes, acesso global a informação, personalização do conteúdo, multimídia e uma cobertura mais intensa de acontecimentos. Porém, esses aspectos se reformulam, reinventam e mudam conforme a mudança da tecnologia e da sociedade, já que ela muda conforme as pessoas também mudam a sua própria cultura. Assim como no cinema, por exemplo, podemos dizer que o jornalismo sempre busca por uma verdade, relatando um caso conforme ele aconteceu, mas fazendo isso do ponto de vista pessoal (no caso de um jornalista, redator, editor com suas crenças pessoais (mesmo que subjetivas), que podem afetar a “verdade” de determinado acontecimento). Ou seja, o jornalista, tal como um cineasta, pode escolher o que mostrar e essa pode ser uma das causas das notícias (e dos sites que divulgam) hiperpartidárias.

Mourão & Robert (2019) alertam para a quantidade de conteúdo que circula na internet ser fabricado, com histórias verdadeiras, mas que apelam ao hiperpartidarismo, que ganha mais espaço nas mídias sociais, tendo um número elevado de compartilhamentos entre as pessoas que creem em algo, mesmo ele não sendo verdadeiro ou superestimado. É especialmente nas mídias sociais que o espaço para desinformar ganha mais força e espaço. Conforme Wardle & Derakhshan (2017) descrevem em seu relatório, preferem rejeitar o termo de “*fake news*” por dois motivos: o processo é mais complexo para ser explicado com esse termo e que ele começou a ser utilizado por líderes políticos mundiais para chamar qualquer notícia/reportagem/investigação (dentre outros) que os desagradem.

Dessa forma, Wardle & Derakhshan (2017) propõem três tipos do que chamam de “desordem informativa”: *misinformation* (informação incorreta que é compartilhada sem a intenção de causar danos), *mal information* (informações verdadeiras, mas que são vazadas com o intuito de causar danos) e o que mais usamos, a *disinformation* (informações falsas compartilhadas deliberadamente para causar danos). Por mais recente que pareça o fenômeno, as técnicas de desinformação não recentes, são especialmente utilizadas em Guerras, utilizando táticas de desinformação para debilitar o adversário e desorientar a opinião pública (DURANDIN, 1993). Para Durandin (1993) o termo desinformação é uma tradução literal vindo do termo russo “*dezinformación*”, denominando campanhas que alguns países capitalistas faziam contra a antiga URSS.

Porém, é com a chegada dos dicionários que o termo começa a ser difundido. Ojegov, em 1949, define a desinformação como a ação de induzir ao erro mediante o uso de informações falsas. Já a “*Grande Encyclopédia soviétique*”, de 1952, define a desinformação como sendo a difusão (pela imprensa, rádio, entre outros) de notícias falsas, com o objetivo de confundir a opinião pública. Essa última definição ainda trazia o viés capitalista da intenção de confundir, visto o clima de guerra existente naquela época. Com o objetivo de conseguir o apoio da opinião pública para suas ações serem bem vistas pela sociedade, as estratégias de guerra também faziam o uso de informações que desinformavam de alguma forma. Mas alguns fatores, como o acesso a internet e a edição rápida de textos, um consumo maior de informação, a velocidade da informação são alguns dos pontos levantados por Wardle & Derakhshan (2017) para que tenhamos a desinformação tal como estamos reconhecendo e convivendo nos dias de hoje.

Acrescentando valor ao termo desinformação (na tradução que optamos em fazer de *disinformation*) Jack (2017) afirma que as plataformas digitais impulsionam a disseminação de informações problemáticas, como exemplo os sites sem fontes, histórias inverificáveis e fabricadas. Colaborando para essa escolha, Fallis (2015) considera que a desinformação incluiria propaganda enganosa, propaganda governamental, fotografias adulteradas, documentos falsificados, mapas falsos, fraudes na internet, sites falsos e até mesmo modificações erradas feitas em sites colaborativos, como a Wikipédia. Com isso, o autor define a *disinformation* como uma informação enganosa que tem como principal função enganar.

Assim entendemos, podemos considerar que o termo “*fake news*” está inserido dentro de um possível “gênero” desinformação (*disinformation*). Tandoc, Lim & Ling (2017), fazem um apanhado geral de como o termo “*fake news*” foi utilizado na academia entre os anos de 2003 e 2017. Primeiro os autores atentam para a importância que a digitalização teve na forma de produção e consumo de notícias: de um lado uma informação pode ser postada no Twitter (via link ou mesmo texto) e ser compartilhada pelo usuário que pode, por sua vez, fazer essa mensagem recircular (ZAGO, 2011), um ambiente digital que oferece uma infinidade de informações, cada vez mais rápido, numa velocidade diferente das mídias que até então conhecíamos. Tandoc et. al. lembram da presença de fontes sociais próximas, já que “receber informações de fontes sociais próximas pode ajudar a legitimar a veracidade das informações compartilhadas

nas redes sociais. No entanto, os usuários raramente verificam as informações que compartilham”<sup>2</sup> (TANDOC, LIM & LING, 2017, p. 139), ou seja, a confiabilidade de uma informação recebida por uma fonte próxima (um familiar, amigo, até mesmo um líder de opinião conhecido, por exemplo) se torna mais provável em aceitar a mensagem como verdadeira.

Ainda conforme a definição de Wardle & Derakhshan (2017, p. 48), devemos perceber alguns tipos de desinformação que são propostas pelos autores, como:

**Sátira ou paródia:** cada vez mais as pessoas recebem notícias em seus *feeds* de notícias em mídias sociais e pode acontecer a confusão quando uma informação não é bem compreendida (no Brasil, por exemplo, o Sensacionalista<sup>3</sup>).

**Conexão falsa:** acontece quando um título, imagem ou legenda não condizem com o conteúdo, são os chamados *clickbait*, feitos para atrair leitores, especialmente quando compartilhados em mídias sociais, usando elementos visuais e legendas que não são apoiadas pelo texto em si.

**Conteúdo enganoso:** quando há o uso enganoso de informações para enquadrar determinados problemas ou indivíduos de alguma forma. Exemplos podem ser mostrados como montagem ou corte de imagens.

**Contexto falso:** nisso determinada informação é retirada de contexto, circulando fora do seu contexto original. Um exemplo disso são fotos de desastres naturais ou tragédias humanas.

**Conteúdo impostor:** quando um conteúdo circula com logotipo de algum veículo de comunicação ou assinatura de alguma pessoa, quando nenhuma dessas fez, realmente, o conteúdo.

**Conteúdo manipulado:** quando um conteúdo que é genuíno é manipulado para enganar.

**Conteúdo fabricado:** geralmente em formato de texto, um conteúdo é falso. Exemplos são os “sites de notícias”, mas pode ser refletida em gráficos criados com dados incorretos.

---

<sup>2</sup> Tradução do autor para: Receiving information from socially proximate sources can help to legitimate the veracity of information that is shared on social networks. However, users seldom verify the information that they share.

<sup>3</sup> Disponível em: <https://www.sensacionalista.com.br/> Acesso em: 20 nov. 2020.

Os estudos de Tandoc (2017) também caminham no mesmo sentido de definição do que seriam *fake news*, sendo que elas seriam uma forma de desinformação:

**Sátira de Notícias** – programas de notícias falsas que utilizam do exagero e humor para apresentar informações diárias. Oferecem primeiro o entretenimento. Esses programas podem moldar o discurso político, opiniões e confiança política. Conteúdo se baseia em fatos reais, mas assumindo o humor, exibindo estilos exagerados, reportagens falsas, risadas... Deixa claro que seu papel não é jornalismo.

**Paródia de Notícias** – partilha muito do que foi definido como sátira. Apresentação imita meios de comunicação. Diferença reside em fatos não factuais. O conteúdo é fabricado.

**Fabricação de Notícias** – que não tem base factual. Não há o entendimento que o conteúdo é falso. O produtor busca *misinforming* (desinformar - missinformin), publicados em sites, blogs ou mídia social. Para ser bem sucedida, o item deve basear-se em memes ou parcialidades pré-existentes. Publicada sob “um olhar” de autenticidade, aderindo estilos e apresentações de notícias. Sites que imitam organizações jornalísticas. O conteúdo fabricado depende da tensão social existente. Pessoas com diferenças políticas, sectárias, raciais ou culturas estão mais vulneráveis às notícias falsas. Bots para as publicações parecerem amplamente divulgadas. Autor da fake news está motivado por razões econômicas ou políticas.

**Manipulação de fotos** – “fake news” foi utilizada para descrever a manipulação de imagens ou vídeos para criar uma narrativa falsa. Fotos retiradas de contexto e de período. Apropriação indébita (exemplo: pessoas indo para manifesto pró-Trump).

**Publicidade e Relações Públicas** – “fake news” foram utilizadas para descrever materiais publicitários sob forma de reportagens genuínas. Cita o caso de vídeos produzidos por relações públicas – mensagens persuasivas com aparência de jornalismo. Fake news: um título que não corresponde com o conteúdo que é clicado, apostando em notícias para atrair a atenção. Pode semear o ódio por algo que nem aconteceu.

**Propaganda** – notícias criadas por organizações políticas para influenciar o público. Narrativas estratégicas, moldar a percepção e ações do público. Difere da publicidade porque baseia-se em fatos, mas inclui preconceitos que promovam um lado, objetivo de persuadir e não informar.

Vimos alguns tipos de desinformação: texto (como forma de fake news), imagens modificadas ou tiradas de contexto, assim como vídeos da mesma forma, mas o alerta atual está para o que especialistas chamam de *deepfakes*. Tandoc (2019), pesquisador reconhecido por definir o objeto *fake news* (2017; 2019) (conforme já relatado neste artigo), alerta para esse novo fenômeno que é, basicamente: “notícias falsas com falsificação profunda, que se referem ao uso de inteligência artificial para alterar digitalmente arquivos de vídeo e áudio<sup>4</sup>” (TANDOC, 2019, p. 2). Ou seja, passamos por mais do que vídeo tirado do contexto, antigos sendo postados com data atual, e começamos a acompanhar falsificações que podem manchar reputações, tentar mudar ou moldar a opinião pública, enganando o espectador que, através do viés de confirmação (COSTA, 2017), passa a acreditar (ou mostra uma tendência a isso) em um produto audiovisual que não é verdadeiro.

Comolli (2015), lembra bem o caso do cinema e do logro e assim a relação do espectador com o que passado através de imagem, já que o espectador, ao consentir de bom grado com o logro cinematográfico, estaria ele mesmo predisposto a uma cumplicidade com os personagens logrados (COMOLLI, 2015, p. 187). O autor também fala da questão da verdade e mentira, que não são possíveis definir (como na lógica ou na matemática), já que são relativos e atraídos por disfarces, preferindo o uso de termos como o de falsidade ou de justeza.

Muitos podem dizer que uma imagem vale mais do que mil palavras<sup>5</sup>, mas existem casos que essas imagens precisam ser montadas, que associações devem ser feitas, cifras, datas, para que se eventualmente a imagem possa ser um signo (COMOLLI, 2005, p. 198), ou seja, há exemplos (como em filmes propagandas de guerra) que a busca pretendida pelos criadores é cativar quem vê, como se a imagem fosse verdadeira, mesmo que para isso precisa-se recorrer ao logro ao espectador, para que esse seja a favor ou contra a uma ação ou causa.

O que Comolli (2015) analisa sobre os filmes de guerra, mostra o que hoje acontece com quem utiliza do avanço da tecnologia para lograr quem assiste, com um objetivo também propagandista: fazer com que o espectador compre uma história,

---

<sup>4</sup> Tradução do autor para: And the problem might soon worsen: Technology experts warn about the next generation of fake news, called “deep fakes,” which refer to the use of artificial intelligence to digitally alter video and audio files.

<sup>5</sup> É uma expressão popular de autoria do filósofo chinês Confúcio, utilizada para transmitir a ideia do poder da comunicação através das imagens.

mesmo ela não sendo real, incentivando a repassar isso através de seus compartilhamentos em mídias sociais ou mesmo através de mensagens por aplicativo de mensagens, que podem propagar desinformação com mais velocidade e com maior dificuldade de verificar o que é repassado.

### O que são *deepfakes*?

Em 2017 Edson Tandoc busca em artigos a definição que era utilizada até então para o termo “*fake news*”, fazendo uma tipologia dos tipos que eram utilizados até então, chegando na seguinte tipologia: sátira; paródia; fabricação, manipulação, publicidade e propaganda, todas elas baseadas em duas dimensões: o nível de factualidade e de engano/fraude. O autor deixa claro que essa tipologia foi feita na análise dos artigos produzidos até então, não concordando que sátira e paródia seriam conteúdos falsos. Porém, temos exemplos como O Sensacionalista<sup>6</sup>, que faz com que o humor seja usado para informar algo verdadeiro, factual, mas que a justificativa tende ao humor, podendo confundir o leitor. E é o mesmo autor que atenta para o começo do uso de *deepfakes*.

Em estudo semelhante ao realizado por Tandoc, Westerlund (2019) analisa 84 artigos sobre o tema das *deepfakes*, com o objetivo de examinar o que elas são, quem produz e benefícios e ameaças possíveis com essa nova possibilidade. É importante destacar, segundo o estudo, que o YouTube se tornou o segundo espaço onde pessoas buscam por notícias, perdendo apenas para o Facebook. Segundo a autora:

Avanços tecnológicos recentes tornaram mais fácil criar o que agora é chamado de “*deepfakes*”, vídeos hiper-realistas usando trocas de rosto que deixam poucos vestígios de manipulação (Chawla, 2019). *Deepfakes* são o produto de aplicativos de inteligência artificial (IA) que mesclam, combinam, substituem e sobrepõem imagens e vídeos para criar vídeos falsos que parecem autênticos (Maras & Alexandrou, 2018). A tecnologia *Deepfake* pode gerar, por exemplo, um vídeo humorístico, pornográfico ou político de uma pessoa dizendo algo, sem o consentimento da pessoa cuja imagem e voz estão envolvidas (Day, 2018; Fletcher, 2018)<sup>7</sup> (WESTERLUND, 2019, p. 39).

<sup>6</sup> Disponível pelo link: <https://www.sensacionalista.com.br/> Acesso em: 20 nov. 2020.

<sup>7</sup> Tradução do autor para: Recent technological advancements have made it easy to create what are now called “*deepfakes*”, hyper-realistic videos using face swaps that leave little trace of manipulation (Chawla, 2019). *Deepfakes* are the product of artificial intelligence (AI) applications that merge,

Percebe-se o uso variado dessa nova tecnologia, especialmente utilizada para o humor. Mas ao passo que essa tecnologia avança, novas áreas podem ser atingidas, assim como o humor pode ser usado pelo criador, mas chegar ao espectador final como algo verdadeiro, visto que por muitas vezes não há uma forma de deixar claro que determinado conteúdo é somente direcionado ao humor (ou por vezes não se quer deixar isso evidente). O que mais colabora para a disseminação deste formato de desinformação é a rapidez que um vídeo pode ser compartilhado nas mídias sociais, fazendo com que as *deepfakes* possam ser mais uma ferramenta eficiente para desinformar.

Westerlund (2019) alerta que o apanhado que fez demonstra que as *deepfakes* podem se tornar uma grande ameaça à sociedade, ao sistema político e as empresas, podem interferir em eleições, prejudicar a confiança do cidadão nas informações e nas autoridades e moldar a opinião pública com um fato inverídico.

### **O humor como principal fator de *deepfakes***

Um bom exemplo que temos, no Brasil, de *deepfake* é o perfil de Bruno Sartori<sup>8</sup> no Instagram. Segundo Sartori, que se auto intitula de “Jornalista, Deepfaker”, em oficina realizada em julho de 2019<sup>9</sup>, explica o uso de troca de rosto feita através de dublês, vídeos e voz para a criação de memes e até mesmo *deepfakes* feitas em tempo real, durante shows de comediantes.

---

combine, replace, and superimpose images and video clips to create fake videos that appear authentic (Maras & Alexandrou, 2018). Deepfake technology can generate, for example, a humorous, pornographic, or political video of a person saying anything, without the consent of the person whose image and voice is involved (Day, 2018; Fletcher, 2018).

<sup>8</sup> Pelo link: <https://gizmodo.uol.com.br/bruno-sartori-cria-deepfakes-bolsonaro-moro/> Acesso em: 20 nov. 2020.

<sup>9</sup> Disponível em: <https://lab.witness.org/wp-content/uploads/sites/29/2019/10/WITNESS-Deepfakes-Prepara-Se-Agora-PT-Updated.pdf> Acesso em: 20 nov. 2020.

Figura 1: Captura de tela do perfil de Bruno Sartori



Fonte: o autor

Como vimos, a sátira e o humor estavam dentre as categorias que apareciam entre formas de desinformar (TANDOC 2017; 2019; WARDLE & DERAKHSHAN, 2017). Isso ainda pode continuar e *deepfakes* utilizadas pelo humor, mas com personagens políticos, podem fazer com que aquele receptor mais desatento passe a acreditar algo que não é real, visto que muitos vídeos, conforme o exemplo abaixo, trazem assuntos factuais (um dos requisitos observados nas formas de desinformar e nas *fake news*), o que pode fazer com que uma peça de humor possa desinformar uma parcela da população.

Figura 2: Captura de tela de vídeo publicado em 6 de julho de 2020



Fonte: o autor

Uma das formas de evitar que vídeos com modificações profundas possam ser utilizados apenas para o humor é rotulando o conteúdo (visivelmente) como humorístico, ou seja, utilizando de uma identificação, para que fique claro a sua intenção, para quando compartilhado fora do local original, ele possa ser identificado (VIZOSO et. al, 2019).

Mas mesmo com todas essas formas, as *deepfakes* podem se popularizar no viés de uma sociedade polarizada que, mesmo sabendo do uso disso para o humor, pode

utilizar para outras formas. Esse risco se torna mais realidade quando os personagens do vídeo modificado passam a ser políticos e representantes de poderes. Vídeos desses, tirados de contexto e não identificáveis, fazem com que as pessoas desconfiem do logro, mas já que o conteúdo vai ao encontro do que se pensa, passa a acreditar, a disseminar e a (provavelmente) desinformar.

É importante reafirmarmos os efeitos que podem ser gerados por uma desinformação. O relato histórico que fizemos dos termos nos mostra que desinformação não é um fenômeno exclusivo na internet e nem nasce com a popularização das mídias sociais pelo mundo. Mas a evolução (tanto técnica/tecnológica e da sociedade atual) traz modificações. Toda a vida, hoje, pode se transferir a uma plataforma online: as fotos que era reveladas e colocadas em quadros, hoje ocupam um mural no Instagram. As amizades que se formavam em festas, locais de trabalho e outros, foram para o Facebook. É o que Pierra Levy (2009) chama de virtualização da vida. Não se nega o real, mas a vida passa a também ser uma experiência virtual.

Nesses espaços virtuais há espaço para transpormos a nossa vida. As conversas, os boatos de boca a boca também são transferidos para uma esfera online. Porém, uma conversa entre duas pessoas, por exemplo, pode demorar um tempo muito grande para chegar a um número maior de pessoas (quando dependem de pessoas comuns para espalharem alguma informação). Esse tempo, nas mídias sociais, pode ser muito reduzido, sendo que uma informação na qual queremos que seja espalhada, pode chegar a um número maior de pessoas do que uma conversa presencial particular.

### **Considerações finais**

A maioria das pessoas tem a tendência de espalhar aquilo (no caso conteúdo falso) que acredita ou que recebe de pessoas em que confia e esse fato é um campo em aberto para a divulgação e circulação de *deepfakes*. Conforme o exemplo que trouxemos nesse artigo, no Brasil (até a presente data de novembro de 2020), as modificações profundas de imagens acontecem na forma de humor ou sátira, duas características que aparecem bastante quando falamos em *fake news*. Porém, essa mesma classificação de humor e sátira são questionadas pelo próprio Tandoc (2019), visto que esses materiais (com o intuito apenas de divertir) são claramente destinada a um propósito: o humor. Mas mesmo com alguma indicação clara no vídeo (uma marca d'água, por exemplo), o

conteúdo que tinha como o intuito apenas divertir (ou uma paródia de algum acontecimento) pode ser direcionado para algum sentido pejorativo, especialmente quando esses vídeos envolvem políticos ou governantes, que foi o exemplo trazido neste artigo.

Porém, ao redor do mundo tem se tornado mais fácil fazer vídeos com pequenas modificações (como filtros em plataformas de mídia social) ou mesmo com estruturas em sites que permite fazer uma modificação profunda em vídeos. Citado por Kehdy et. al. (2018) temos o exemplo do site para o público adulto<sup>10</sup> que permite que o usuário utilize imagens de terceiros e coloque os rostos que bem entender, fazendo com que em poucos minutos, o rosto de uma pessoa que nunca tenha participado de vídeo desse estilo, apareça em um. Outro exemplo que viraliza de tempos em tempos no Brasil é o FaceApp, um aplicativo simples que consegue fazer modificações fáceis e intuitivas que faz com que adicione cabelo, olhos diferentes e até mesmo envelhecer determinada imagem de uma pessoa.

Percebemos o potencial que esse tipo de modificação profunda de vídeos e imagens pode fazer. Alguns casos emblemáticos, como do então candidato ao Governo do Estado de São Paulo João Dória Júnior<sup>11</sup>, que teve um suposto vídeo íntimo divulgado em meio a sua campanha eleitoral do ano de 2018. O foi desmentido pelo próprio e por sua assessoria, mas ganhou boa parte do noticiário daquela semana. Casos como esse buscam aquilo que a desinformação deseja: destruir ou arranhar a imagem de alguém ou de alguma instituição. Ainda mais quando esse tipo de possível manipulação acontece em meio a uma eleição.

A legislação brasileira ainda caminha para fins de evitar as possíveis *fake news* e busca formas para coibir a desinformação. Mas é preciso pensar não somente nas opções jurídicas, mas sim na solução e na busca de softwares livres para que possamos identificar conteúdos modificados, já que por outro lado eles também contribuem para a confecção de *deepfakes* (KEHDY ET. AL., 2018). Ou seja, a filosofia de software livre permite, hoje, a criação de *deepfakes*, mas deve vir a colaborar com a população e com

---

<sup>10</sup> Disponível pelo link: [adultdeepfakes.com](http://adultdeepfakes.com). Acesso em: 20 nov. 2020.

<sup>11</sup> Disponível em: [https://br.noticias.yahoo.com/internet-veio-abaixo-com-o-suposto-video-intimo-doria-143030578.html?guccounter=1&guce\\_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce\\_referrer\\_sig=AQAAADvA--uXq\\_maeJo725KgSi1iQ3OcMUYSq5BGNcE4XmVLXPx6S66rUVJcVVZtwuj9w37jPicYKAs0DAzvFn\\_MwjcAp3IzEJTFrYwZERX-0haNHYMUOZgpUwIupyaTF6Pj5-DV8F5CT9UyOaPmLPhZI2jSe6PunUrxpnwkpG4p3tXp](https://br.noticias.yahoo.com/internet-veio-abaixo-com-o-suposto-video-intimo-doria-143030578.html?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAADvA--uXq_maeJo725KgSi1iQ3OcMUYSq5BGNcE4XmVLXPx6S66rUVJcVVZtwuj9w37jPicYKAs0DAzvFn_MwjcAp3IzEJTFrYwZERX-0haNHYMUOZgpUwIupyaTF6Pj5-DV8F5CT9UyOaPmLPhZI2jSe6PunUrxpnwkpG4p3tXp) Acesso em: 20 nov. 2020.

a sociedade na busca pela verdade, por uma democracia e por uma comunicação que seja transparente com quem recebe.

Buscamos mostrar um conjunto de referências que demonstrem como a modificações profundas de imagens podem trazer uma confusão e gerar um novo tipo de desinformação. Dentre vários tipos de *disinformation* (desinformação) como *fake news*, por exemplo, as *deepfakes* ganham força por apelar ao visual e toda a sociedade estar acostumada a este aspecto de dar relevância naquilo que se vê. O exemplo de João Dória mostra o quão nocivo essa modificação pode ser para uma pessoa ou alguma instituição. Precisamos de esforços (como a criação de software livres e legislação) para que esses tipos de vídeos fiquem apenas no humor e não penetrem em outras esferas da sociedade, vindo a prejudicar a sociedade com desinformação.

Ainda sim, percebemos a limitação da pesquisa de *deepfakes* no Brasil pelo fato de serem usadas, em sua maioria, para humor, paródia ou sátira de alguma situação, que são compartilhadas em mídias sociais. Porém, esse conteúdo pode circular em grupos fechados ou de mensagens, que se torna um limitando em saber o que está circulando nessas plataformas.

## Referências

CASTELLS, Manuel. **A sociedade em rede**. São Paulo: Paz e Terra, 1999. v. 1.

COMOLLI, Jean-Louis. **Pensar o cinema-imagem, ética e filosofia**. São Paulo: Cosac Naify, 2015, p. 165-203.

COSTA, Daniel. Ensaio sobre vieses cognitivos no processo de tomada de decisão gerencial. Tese (doutorado) – Universidade Federal de Lavras, 2017.

DURANDIN, Guy. **La información, la desinformación y la realidad**. Ediciones Paidós Ibérica, Barcelona, Espanha, 1993.

FALLIS, Don. **What is disinformation?** Library Trends, Volume 63, Number 3, Winter 2015, pp. 401-426.

JACK, Caroline. **Lexicon of lies: terms for problematic information**. Data & Society Research Institute, 2017.

KEHDY ET. AL. **Implicações políticas, sociais e éticas da popularização de deepfakes por meio de softwares livres**. Consciência, UEaDSL, 2018.

LÉVY, Pierre. **O que é virtual?** Tradução Paulo Neves. 2 ed. São Paulo: Editora 34, 2009

MOURÃO, Rachel & ROBERTSON, Craig. **Fake news as discursive integration: an analysis of sites that publish false, misleading, hyperpartisan and sensational information.** *Journalism Studies*, 1–19, 2019.

PAVLIK, John. **El periodismo y los nuevos médios de comunicación.** Paidós, Barcelona, 2005, 351 páginas.

TANDOC, Edson C; LIM, Zheng & LING, Richard. **Defining “Fake News”**, *Digital Journalism*, 137-153, 2017.

TANDOC, Edson. **The facts of fake news: a research review.** *Wiley Connections*, v 13, setembro 2019. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/soc4.12724>> Acesso em: 14 set. 2020.

VIZOSO, Angel et. al. **Humor y desinformación. Espacios de creación de noticias falsas, redes sociales y plataformas de verificación en España.** Conference: Humour and misinformation. Fake news creation sites, social media and verification platforms in SpainAt: Porto, Portugal, 2019.

WARDLE, Claire; DERAKHSTHAN, Hossein. **Information Disorder: toward an interdisciplinary framework for research and policymaking.** Council of Europe, 2017. Disponível em: <<https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c>> Acesso em: 11. set. 2020.

WESTERLUND, Mika. **The Emergence of deepfake technology: a review.** *Technology Innovation Management Review*, v. 9, n. 11. November, 2019.