
A ameaça da *deepfake* na sociedade contemporânea: uma reflexão à luz de 'Joan é Péssima' de *Black Mirror*¹

Natã FREITAS²

Talita Souza MAGNOLO³

Universidade Federal de Juiz de Fora, Minas Gerais

RESUMO: O artigo analisa o episódio "Joan é Péssima" da série *Black Mirror*, destacando a mensagem central sobre os perigos da tecnologia, especificamente os *deepfakes*. A série *Black Mirror* é conhecida por retratar futuros distópicos onde a tecnologia tem consequências devastadoras. No episódio, a protagonista Joan Clayton é vítima de um *deepfake* que destrói sua vida pessoal e profissional. A narrativa alerta para a proliferação de que podem difamar, extorquir e manipular informações. Utilizando a metodologia da Análise da Materialidade Audiovisual (Coutinho, 2016), a análise pretende explorar os significados culturais e sociais do episódio, destacando as preocupações contemporâneas com a privacidade e a manipulação da informação.

PALAVRAS-CHAVE: *Deepfake*; *Black Mirror*; Inteligência Artificial; Análise da Materialidade Audiovisual.

RESUMO EXPANDIDO:

O avanço implacável da tecnologia nos últimos anos é inegável (Boden, 2020). Seus benefícios, incontáveis, transformaram diversos aspectos da vida humana, no entanto, essa mesma força propulsora, quando utilizada de forma indevida ou sem discernimento, pode se manifestar como um monstro devorador, capaz de causar danos irreparáveis. Essa é a mensagem central do episódio "Joan é Péssima" – em inglês *Joan Is Awful* - da série *Black Mirror*, da plataforma de streaming Netflix, que nos convida a uma reflexão profunda sobre os perigos nas sombras do progresso tecnológico.

Black Mirror, série aclamada pela crítica, não se contenta em apenas retratar o futuro, ela o esculpe, moldando-o em um cenário distópico onde a tecnologia reina suprema, muitas vezes com consequências devastadoras para a humanidade (Nuntiare, meio digital, 2024). "Joan é Péssima" é um exemplo magistral dessa abordagem, servindo

¹ Trabalho apresentado no GP Ficção Televisiva Seriada do XIX Encontro dos Grupos de Pesquisas em Comunicação, evento componente do 47º Congresso Brasileiro de Ciências da Comunicação.

² Graduando em Rádio, TV, Internet na Faculdade de Comunicação da UFJF. Bolsista de Extensão do Projeto Memória. E-mail: Nata.freitas@estudante.ufjf.br

³ Professora Substituta da Faculdade de Comunicação da UFJF. Coordenadora do Projeto de Extensão Memória. Vice-líder do Grupo de Pesquisa COMCIME. E-mail: talita.magnolo@yahoo.com.br

como um alerta perturbador sobre os perigos dos deepfakes, vídeos manipulados por Inteligência Artificial para mostrar ou fazer alguém dizer algo que nunca disse ou fez.

A Inteligência Artificial (IA) é um ramo da ciência da computação que lida com a criação de sistemas que podem realizar tarefas que normalmente requerem inteligência humana (Boden, 2020). Isso inclui aprendizado, raciocínio, resolução de problemas, percepção e linguagem natural. O Aprendizado de Máquina (AM) é o campo de estudo que dá aos computadores a capacidade de aprender a partir de dados, sem serem explicitamente programados. De acordo com Gabriel (2022), existem três tipos principais de aprendizado de máquina: o aprendizado supervisionado, em que são inseridos os dados de entrada e saída, tornando a resposta mais assertiva; o aprendizado não supervisionado, em que são inseridos apenas os dados de entrada; o aprendizado por reforço, em que a máquina toma determinadas decisões baseada no ambiente e contextos fornecidos.

No episódio que será analisado, a protagonista Joan Clayton se vê vítima de um deepfake cruel e humilhante, que destrói sua reputação, carreira e relacionamentos. A narrativa, embora ficcional, nos confronta com uma realidade cada vez mais próxima: a proliferação de deepfakes, com seu potencial para causar danos irreparáveis à vida das pessoas. De acordo com Santaella (2018), a manipulação da informação através da Inteligência Artificial se torna uma arma letal, capaz de difamar, extorquir e até mesmo influenciar eleições.

Recentemente, nos deparamos com vários casos polêmicos, como, por exemplo, a imagem do Papa Francisco usando um grande casaco branco inundou a internet em março de 2023, justamente porque muitas pessoas não sabiam que ela foi criada com inteligência artificial. Na época, ferramentas alimentadas por IA, como o *MidJourney*, que permite ao usuário criar imagens realistas usando apenas comandos de texto, eram novidade. (Forbes, meio digital, 2024).

O Colégio Santo Agostinho, da Barra da Tijuca, no Rio de Janeiro, enviou um comunicado aos pais sobre os alunos estarem utilizando a IA para criar montagens de fotos íntimas, no total foram 20 alunas afetadas. (Revista Oeste, meio digital, 2024). Além disso, o caso do funcionário de multinacional em Hong Kong, vítima de um golpe milionário utilizando *deepfakes* e inteligência artificial, expõe a face sombria dessa tecnologia: a manipulação e o engano. Essa narrativa serve como um alerta para os perigos dessa ferramenta, que pode ser usada para fins maliciosos, causando danos financeiros (CNN, meio digital, 2024).

Embora o episódio "Joan é Péssima" se desenrole em um futuro distópico, as conexões com a nossa realidade são inquietantes, como a rápida evolução da tecnologia da informação, a crescente dependência das mídias sociais e a falta de senso crítico diante da avalanche de informações online criam um terreno fértil para a proliferação de deepfakes, conforme apontado por Faustino e Lippold (2023).

Os *deepfakes* são frequentemente utilizados para fins humorísticos nas redes sociais, mas, podem, também, ter consequências graves para as pessoas retratadas nessas imagens e vídeos falsos gerados por inteligência artificial. Uma pesquisa recente do grupo *DeepTrace Labs* (CanalTech, meio digital, 2024) mostrou que cerca de 96% dos *deepfakes* são pornográficos ou imagens de sexo não-consensual. Muitos são de celebridades e pessoas públicas, mas há também de indivíduos privados.

De forma resumida, no episódio "Joan é Péssima", somos apresentados à vida aparentemente normal de Joan, uma mulher comum com seus problemas e alegrias do dia a dia. No entanto, essa realidade se transforma em um pesadelo quando ela se depara com a criação de um *deepfake* extremamente realista que retrata sua vida em detalhes, exagerando eventos e expondo segredos. A partir desse ponto, acompanhamos os impactos dessa tecnologia na vida de Joan, que perde seu emprego e sofre danos irreparáveis à sua reputação e saúde mental.

Em certo ponto do episódio, Joan procura sua advogada com o objetivo de processar a Strawberry Corporation, a empresa de streaming responsável pela produção da série sobre sua vida. Ela fica surpresa ao descobrir que a empresa tem o direito de fazê-lo, uma vez que concordou com os termos e condições ao criar sua conta no serviço de streaming.

Faustino e Lippold (2023) abordam a questão do consentimento na era digital, criticando a maneira como as grandes empresas de tecnologia coletam e usam os dados de seus usuários. Os autores afirmam que, frequentemente, os termos de consentimento são extensos, complexos e difíceis de entender, resultando em uma situação de "colonialismo digital", onde os usuários são expostos a práticas exploratórias de dados sem pleno conhecimento ou consenso.

Ao ignorar os termos de uso, abrimos mão do controle sobre nossos dados pessoais e privacidade. Esses termos podem conter cláusulas que concedem às empresas o direito de coletar, armazenar e utilizar nossas informações para diversos fins, desde publicidade

direcionada até a criação de perfis comportamentais. Essa falta de transparência e controle sobre nossos dados representa um risco significativo à nossa privacidade e segurança.

Imagem 1 – Extrato do episódio “Joan é Péssima”



Fonte: Netflix

Para a análise, utilizaremos a metodologia da Análise da Materialidade Audiovisual:

É nessa perspectiva que no âmbito das pesquisas realizadas no Laboratório de Jornalismo e Narrativas Audiovisuais (CNPq-UFJF) tem-se buscado o desenvolvimento de um método denominado de Análise da Materialidade Audiovisual, que tomaria como objeto de avaliação a unidade texto + som + imagem + tempo + edição (Coutinho, 2016).

A metodologia propõe que sejam reconhecidos os sentidos propostos pelo produto audiovisual, antes de iniciar o processo de análise. Sendo assim, é possível afirmar que o episódio reflete as preocupações e debates da sociedade contemporânea sobre o *deepfake*, seus impactos éticos e sociais, bem como as potenciais ameaças à privacidade.

A partir disso, será possível uma análise que compreenda todos os aspectos do produto audiovisual, entre eles, a produção, o canal de disseminação, seus diálogos com a realidade, a construção dos personagens, músicas e edição. Acreditamos que este método tornará possível uma compreensão profunda do contexto social, cultural e político em que ele foi produzido e dos significados mais amplos que ele carrega. Essa abordagem revelará as preocupações e angústias da sociedade contemporânea em relação aos perigos

e desafios da IA, especialmente no que tange à manipulação da informação e à construção de realidades alternativas.

A facilidade e a praticidade da tecnologia moderna podem nos levar a uma confiança cega, ignorando os riscos subjacentes. A partir do contexto apresentado, é possível elaborar uma analogia: assim como nossos ancestrais confiaram no fogo sem precauções adequadas, corremos o risco de cair nas armadilhas da era digital. Termos contratuais impressos ou enviados por e-mail são minuciosamente analisados antes da assinatura, mas os termos de uso online, incluindo os "cookies", são frequentemente ignorados sem a devida atenção (R7 Notícias, meio digital, 2024).

A analogia do fogo, criada para ser comparada à IA, pode ser expandida para outras tecnologias: a roda facilitou o transporte, mas também gerou acidentes; a imprensa democratizou o acesso à informação, mas também propagou desinformação; e a inteligência artificial promete revolucionar diversos setores, mas também levanta preocupações éticas e de segurança.

Esperamos, com a análise, poder demonstrar que o episódio nos leva a refletir sobre os perigos dos *deepfakes*, que podem ser utilizados para criar imagens e vídeos falsos com extrema precisão, manipulando a realidade e causando danos à vida das pessoas. Além disso, acreditamos que a análise comprovará que "Joan é Péssima" não se limita a apresentar um futuro distópico, mas serve como um espelho para a nossa realidade, convidando-nos a questionar o uso responsável da tecnologia, a importância da responsabilidade individual e a construção de uma sociedade mais consciente e crítica em relação à manipulação da informação e à invasão da privacidade.

REFERÊNCIAS

BODEN, Margareth. **Inteligência Artificial: Uma brevíssima introdução**. São Paulo: unesp, 2020

CANALTECH. 96% dos deepfakes são pornográficos, segundo relatório. Disponível em: <https://canaltech.com.br/internet/96-dos-deepfakes-sao-pornograficos-segundo-relatorio-151749/>. Acesso em: 10 jun. 2024.

CNN BRASIL. **Funcionário de multinacional paga US\$ 25 mi a golpista que usou "deepfake" para simular reunião**. Disponível em: <https://www.cnnbrasil.com.br/internacional/funcionario-de-multinacional-paga-us-25-mi-a-golpista-que-usou-deepfake-para-simular-reuniao/> Acesso em: 10 Jun. 2024.

COUTINHO, Iluska. **O telejornalismo narrado nas pesquisas e a busca por cientificidade: A análise da materialidade audiovisual como método possível.** XXXIX Congresso Brasileiro de Ciências da Comunicação – São Paulo, 2016.

ESTADÃO VERIFICA. **Deepfakes atingem principalmente mulheres, alerta especialista.** Disponível em: <https://www.estadao.com.br/estadao-verifica/deepfakes-atingem-principalmente-mulheres-alerta-especialista/>. Acesso em: 10 Jun. 2024..

FAUSTINO, Deivison; LIPPOLD, Walter. **Colonialismo digital: por uma crítica hacker-fanoniana.** São Paulo: Boitempo, 2023.

FORBES. **Papa Francisco alerta sobre perigos da IA após foto viral** Disponível em: <https://forbes.com.br/forbes-tech/2024/01/papa-francisco-alerta-sobre-perigos-da-ia-apos-foto-viral/> Acesso em: 10 Jun. 2024.

GABRIEL, Martha. **Inteligência Artificial: do zero ao metaverso.** Barueri: atlas, 2022.

NUNTIARE. **Isso é muito Black Mirror... ou não?** Disponível em: <https://www2.uepg.br/nuntiare/isso-e-muito-black-mirror-ou-nao/>. Acesso em: 10 jun. 2024.

R7 NOTÍCIAS **Metade dos brasileiros aceita cookies digitais sem saber o que são diz pesquisa** <https://noticias.r7.com/tecnologia-e-ciencia/metade-dos-brasileiros-aceita-cookies-digitais-sem-saber-o-que-sao-diz-pesquisa-13072022/> Acesso em: 10 Jun. 2024

REVISTA OESTE **Alunos de colégio do Rio usam IA para criar imagens íntimas de meninas** Disponível em: <https://revistaoeste.com/brasil/alunos-de-colegio-do-rio-usam-ia-para-criar-imagens-intimas-de-meninas/> Acesso em: 10 Jun. 2024.

SANTAELLA, Lúcia. **A inteligência é inteligente?.** In: A inteligência e suas múltiplas faces. São Paulo: Cortez, 2016.