

PROPOSTAS DE UMA MORATÓRIA GLOBAL DA PESCA DE ARRASTO DE FUNDO: CENÁRIOS E POSSIBILIDADES JURÍDICAS

Gustavo Silveira Borges 

FFabio Jeremias de Souza 

Contextualização: O avanço da tecnologia e o aumento dos usuários adeptos nas mídias sociais tem ampliado o efeito nocivo da desinformação para a democracia. Nesse sentido, diversos estudos estão sendo desenvolvidos pelo mundo para buscar o combate à desinformação, com destaque para o Relatório da ONU sobre desinformação e liberdade de opinião e expressão de 2021.

Objetivo: O artigo tem por objetivo geral estudar em que medida as diretrizes do Relatório da ONU sobre desinformação e liberdade de opinião e expressão (2021 - A/HRC/47/25) auxiliarão no combate à desinformação nas mídias sociais. Num primeiro momento o estudo apresenta uma análise conceitual sobre a desinformação, liberdade de opinião e liberdade de expressão. Após, buscou-se apresentar como funciona a lógica da moderação de conteúdo sobre desinformação pelas mídias sociais para, ao final, apresentar as diretrizes e recomendações da ONU sobre o tema.

Metodologia: Utilizou-se o método de abordagem dedutivo, eis que se irá partir de premissas gerais a fim de chegar a uma conclusão particular, bem como pesquisa bibliográfica.

Resultados: Conclui-se que as diretrizes do Relatório da ONU sobre desinformação e liberdade de expressão oferece mecanismos e recomendações que nortearão o combate à desinformação.

Palavras-chave: Mídias Sociais; Desinformação; Moderação de Conteúdo; Liberdade de Expressão; ONU.

DISINFORMATION, FREEDOM OPINION AND EXPRESSION AND CONTENT MODERATION ON SOCIAL MEDIA

Contextualization: The advancement of technology and the increase of adept users in social media have amplified the harmful effect of disinformation for democracy. In this sense, several studies are being developed around the world to seek to combat disinformation, with emphasis on the 2021 UN Report on Disinformation and Freedom of Opinion and Expression.

Objective: The general objective of the article is to study the extent to which the guidelines of the UN Report on Disinformation and Freedom of Opinion and Expression (2021 - A/HRC/47/25) will help to combat disinformation on social media. At first, the study presents a conceptual analysis of disinformation, freedom of opinion and freedom of expression. Afterwards, we sought to present how the logic of content moderation on disinformation through social media works to, in the end, present the UN guidelines and recommendations on the subject.

Methodology: The deductive method of approach was used, since it will start from general premises in order to reach a particular conclusion, as well as bibliographic research.

Results: It is concluded that the guidelines of the UN Report on disinformation and freedom of expression offer mechanisms and recommendations that will guide the fight against disinformation.

Keywords: Social Media; Disinformation; Content Moderation; Freedom of Expression; UN.

DESINFORMACIÓN, LIBERTAD DE OPINIÓN Y EXPRESIÓN Y MODERACIÓN DE CONTENIDO EN REDES SOCIALES

Contextualización: El avance de la tecnología y el aumento de usuarios en las redes sociales han amplificado el efecto nocivo de la desinformación para la democracia. En este sentido, se están desarrollando varios estudios a nivel mundial para buscar combatir la desinformación, con énfasis en el Informe de la ONU sobre Desinformación y Libertad de Opinión y Expresión 2021.

Objetivo: El objetivo general del artículo es estudiar en qué medida los lineamientos del Informe de la ONU sobre Desinformación y Libertad de Opinión y Expresión (2021 - A/HRC/47/25) ayudarán a combatir la desinformación en las redes sociales. En un primer momento, el estudio presenta un análisis conceptual de la desinformación, la libertad de opinión y la libertad de expresión. Posteriormente, buscamos presentar cómo funciona la lógica de moderación de contenidos sobre la desinformación a través de las redes sociales para, al final, presentar los lineamientos y recomendaciones de la ONU sobre el tema.

Metodología: Se utilizó el método de enfoque deductivo, ya que se partirá de premisas generales para llegar a una conclusión particular, así como la investigación bibliográfica.

Resultados: Se concluye que los lineamientos del Informe de la ONU sobre desinformación y libertad de expresión ofrecen mecanismos y recomendaciones que orientarán la lucha contra la desinformación.

Palabras clave: Redes Sociales; Desinformación; Moderación de contenido; La libertad de expresión; Naciones Unidas.

INTRODUÇÃO

O avanço da tecnologia, o caráter privado das empresas de mídias sociais, a repercussão coletiva da atividade empresarial, a adesão crescente da população mundial nas principais plataformas e, por fim, o fato de que as maiores empresas do setor operam em escala global, faz com que o tema da desinformação ganhe cada vez mais atenção dos Estados, da sociedade civil e da academia.

Em 2017, a ONU elaborou declaração conjunta entre o Relator Especial das Nações Unidas (ONU) sobre Liberdade de Opinião e Expressão, o Representante da Organização para a Segurança e Cooperação na Europa (OSCE) sobre Liberdade de Mídia, o Relator Especial da Organização dos Estados Americanos (OEA) sobre Liberdade de Expressão e a Relatora Especial da Comissão Africana dos Direitos Humanos e dos Povos (CADHP) para a Liberdade de Expressão e Acesso à Informação, oportunidade em que foi enfatizado que a desinformação pode prejudicar uma reputação pessoal, incitar a violência e a discriminação.¹

Naquela oportunidade já se percebia o dilema entre a liberdade de expressão/opinião e o crescente fluxo de desinformação, com potencial para prejudicar a democracia e corroer as instituições. Ressaltou o documento que o direito de difundir informação não se limitava às informações corretas, sendo que as proibições de combate à desinformação poderiam violar os padrões internacionais de direitos humanos. Contudo, isso não justifica a divulgação de declarações falsas de forma consciente ou imprudentemente por agentes oficiais.²

Com o aumento gradativo do problema e diante da necessidade de apresentar proposições para o enfrentamento do tema, a Relatora Especial da Organização das Nações Unidas sobre a promoção e proteção do direito à liberdade de opinião e expressão, Irene Khan, após contribuições de 119 organizações da sociedade civil e entidades acadêmicas, 3 organizações internacionais, 3 Estados Membros e 3 empresas, além de reuniões com Estados Membros, empresas de mídia e especialistas, editou o relatório A/HRC/47/25 com diversas conclusões e recomendações.

Assim, o presente trabalho tem por objetivo geral examinar em que medida as diretrizes do Relatório da ONU sobre desinformação e liberdade de opinião e expressão (2021 - A/HRC/47/25) auxiliarão no combate à desinformação nas mídias sociais.

No primeiro capítulo, serão apresentados o conceito de desinformação, de

¹ ORGANIZATION OF AMERICA STATES. Joint declaration of freedom of expression and "Fake News", disinformation and propaganda. 2017. Disponível em <https://www.osce.org/files/f/documents/6/8/302796.pdf>. Acesso em 15 de mai. 2022

² ORGANIZATION OF AMERICA STATES. Joint declaration of freedom of expression and "Fake News", disinformation and propaganda. 2017.

liberdade de opinião e expressão, com o intuito de demonstrar as nuances que tornam desafiadora a análise do tema.

O segundo capítulo trata especificamente da forma com que as mídias sociais tratam a moderação de conteúdo, com um enfoque nos *standards* de direitos humanos e nos padrões da comunidade do *Facebook* sobre desinformação, por ser a maior mídia social em atividade.

O terceiro capítulo apresenta as principais diretrizes do Relatório da Relatora Especial da Organização das Nações Unidas sobre a promoção e proteção do direito à liberdade de opinião e expressão, Irene Khan (A/HRC/47/25) com diversas conclusões e recomendações.

Deste modo, o problema da presente pesquisa consiste em saber: em que medida as diretrizes do Relatório da ONU sobre desinformação e liberdade de opinião e expressão (2021 - A/HRC/47/25) auxiliarão no combate à desinformação nas mídias sociais?

Este artigo se justifica pela atualidade do tema, pelo avanço das novas tecnologias e pela expansão das mídias sociais, considerando que o estudo da desinformação tem recebido atenção pela proliferação coordenada de notícias falsas ou distorcidas, assim como pelo aparente conflito com o direito de opinião e de liberdade de expressão.

A presente pesquisa adotou o método de abordagem dedutivo, eis que partiu de premissas gerais a fim de chegar a uma conclusão particular. Por sua vez, a técnica de pesquisa a bibliográfica, uma vez que se utilizou de livros e artigos científicos já elaborados.

1. DESINFORMAÇÃO NAS MÍDIAS SOCIAIS E LIBERDADE DE EXPRESSÃO E OPINIÃO

A sociedade moderna tem como uma de suas principais características a influência do uso da tecnologia digital, com os benefícios dela decorrentes. No entanto, surge a partir do crescimento desordenado, uma série de consequências, dentre elas, a preocupação com a proteção dos direitos humanos.³

O reflexo das novas tecnologias nas mídias sociais impulsionou o seu desenvolvimento o que, somado à adesão crescente do número de usuários, gerou nos últimos anos uma transformação radical na forma de comunicação e interação entre as pessoas, passando de um modelo público de comunicação de “poucos para muitos”, para uma estrutura de “muitos para muitos”.⁴ O aumento do fluxo de informação fez mudar

³ IZA, Daniela Gonzales, Human Rights In The Digital Era: Challenges and Opportunities From The United Nations Human Rights System. *Jurnal Pengajian Umum Asia Tenggara*. 22 (2021): 1 – 13. 1 p.

⁴ SANDER, Barrie. Democratic Disruption in the Age of Social Media: Between Marketized and Structural Conceptions of Human Rights Law. *The European Journal of International Law*. Vol. 32 no. 1, 2021. p. 163.

profundamente os hábitos de consumo dessas informações, seja pela facilidade de exposição, seja pela democratização dos produtores da informação, assim como pela utilização de algoritmos. O resultado é a redução do tempo dedicado à informação, de forma que os usuários, de um modo geral, acabam por prestar atenção apenas no título das matérias, no seu visual, sem muito apego à fonte e ao conteúdo completo.⁵

Para Aieta⁶, essa sociedade algorítmica significa que decisões importantes, sob o ponto de vista econômico e social, são tomadas por robôs, algoritmos e pela inteligência artificial. Os algoritmos podem ser explicados por sequências finitas de ações executáveis com o objetivo de obter uma solução para um problema, através de fórmulas matemáticas, o que reduz o grande esforço que seria para fazer uma seleção entre os milhares de comentários e notícias selecionadas.

Aliás, o modelo de negócio das mídias é baseado nos algoritmos, em publicidade e no direcionamento dos usuários aos conteúdos extremistas e sensacionalistas, o que em larga medida prejudica o direito de opinião e a liberdade de expressão. Em contrapartida, pode-se dizer que as plataformas de mídia social não parecem ter procurado rever o modelo de negócios projetado para promover conteúdo sensacionalista, manter os usuários conectados às plataformas, fazendo da desinformação uma forma de ganhar dinheiro na era digital.⁷

Assim, de forma paralela ao crescimento das mídias sociais como fonte de informação, houve um aumento de informações distorcidas ou inverídicas, tanto por interesses políticos como financeiros, fenômeno esse designado *fake news*. A desinformação, aliás, é um poderoso instrumento de controle político e social.⁸ Uma vez sendo divulgada em grandes proporções nesse mundo de ritmo frenético, constitui-se numa ameaça para as democracias, diante do impacto negativo na opinião política.⁹

Pelo seu caráter polarizador, além de impedir que as pessoas exerçam de forma absoluta seus direitos humanos, ela corrói a confiança nas instituições.¹⁰ Recentemente,

⁵ STARTS, Janis. Disinformation as a Threat to National Security. 2018. Em **Disinformation and Fake News**. 23-32. Singapore: Palgrave Macmillan. DOI: <https://doi.org/10.1007/978-981-15-5876-4>, p. 25

⁶ AIETA, Vania Siciliano. O Impacto Eleitoral Resultante da Manipulação das Fake News no Universo das Redes Sociais: a Construção da Desinformação. **Revista Interdisciplinar de Direito**. v. 18, n. 1, pp 213-233, jan./jun. 2020. p. 216.

⁷ UNITED NATIONS GENERAL ASSEMBLY. Disinformation and freedom of opinion and expression. Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Irene Khan A/HRC/47/25. Disponível em <https://www.ohchr.org/en/calls-for-input/reports/2021/report-disinformation> Acesso em 15 mai. 2022. p. 14

⁸ AIETA, Vania Siciliano. O Impacto Eleitoral Resultante da Manipulação das Fake News no Universo das Redes Sociais: a Construção da Desinformação. p.218.

⁹ MANYANA, Candice Chirwa and Zimkhitha. The Rise of Fake News: Surveying the Effects of Social Media on Informed Democracy. **The Thinker**. Volume 88, 2021. Disponível em https://journals.uj.ac.za/index.php/The_Thinker/article/view/604/378. Acesso em: 18 mai. 2022

¹⁰ UNITED NATIONS GENERAL ASSEMBLY. Disinformation and freedom of opinion and expression. p. 2

com a pandemia, houve um aumento exponencial da desinformação através das mídias sociais, ao passo em que a narrativa contra o sistema fez diminuir a confiança da população nas instituições públicas.¹¹

Contudo, não há consenso sobre o conceito de desinformação. O Relatório da ONU sobre desinformação e liberdade de opinião e expressão (2021) pontua essas dificuldades: a) não há uma linha clara entre a realidade e a falsidade; b) não há uma linha clara entre a intenção de causar dano ou a ausência dela; c) enquanto informações verídicas podem ser classificadas como “fake news”, as opiniões, crenças, falta de conhecimento e outras formas de expressão, como paródia e sátira, não são facilmente identificadas como verdadeiras ou falsas; d) conteúdos falsos divulgados com intenção de causar danos podem ser compartilhados por terceiros inocentes sem intenção, agregando credibilidade à informação falsa ou distorcida; e) algumas formas de desinformação podem conter incitação ao ódio, discriminação e violência, que são proibidas pelo direito internacional.¹²

No Relatório da ONU, desinformação (*disinformation*) é entendida como informação falsa com intenção de causar sérios danos sociais, e desinformação (*misinformation*) é a disseminação de informação falsa sem o conhecimento e intenção de causar dano.¹³

Já o *Policy Department for Citizen Rights Constitutional Affairs* faz a divisão em três categorias: a) *Misinformation*: Informação falsa, que não possui como objetivo causar lesão; b) *Disinformation*: Divulgação de uma informação falsa de forma deliberada para causar lesão; c) *Malinformation*: divulgação verídica, com o intuito de causar dano.¹⁴ Alguns autores, porém, como no relatório da ONU, dividem apenas em dois grupos, *misinformation* e *disinformation*.¹⁵

Ainda, a desinformação pode ser subdividida em três classes: a) falsidades e rumores distribuídos de forma consciente, podendo esse primeiro grupo fazer parte de uma agenda política com viés ideológico ou uma campanha de desinformação patrocinada pelo Estado; b) falsidades propagadas sem um objetivo político, onde se encaixam as “*misinformations*”, num contraponto às “*disinformations*”, contudo, podendo “viralizar”

¹¹ VÉRITER, Sophie. European Democracy and Counter-Disinformation: Toward a New Paradigm?. **Carnegie Europe**, 2021. Disponível em: <<https://carnegieeurope.eu/2021/12/14/european-democracy-and-counter-disinformation-toward-new-paradigm-pub-85931>>. Acesso em: 24 fev. 2022

¹² UNITED NATIONS GENERAL ASSEMBLY. Disinformation and freedom of opinion and expression. p.3

¹³ UNITED NATIONS GENERAL ASSEMBLY. Disinformation and freedom of opinion and expression. p.3

¹⁴ BARBOSA, Mafalda Miranda. Fake News e Fact-checkers: uma perspectiva jurídico-civilista. **Revista de Direito e Responsabilidade**. Ano 3 – 2021. p. 733-766. p.759

¹⁵ ANG, Benjamin; ANWAR, Nur Diyanah; JAYAKUMAR, Shashi. Disinformation & Fake News: Meanings, Present, Future, 2018. Em **Disinformation and Fake News**. 3-20. Singapore: Palgrave Macmillan. DOI: <https://doi.org/10.1007/978-981-15-5876-4>. p.7

em prejuízo aos indivíduos ou à sociedade; c) falsidades com a finalidade de se obter ganhos financeiros.¹⁶

Starts¹⁷ também faz uma distinção entre a *misinformation* e a *disinformation*, acrescentando alguns elementos. A primeira classe, *misinformation*, seria informação falsa ou enganosa, mas não maliciosa, criada por uma pessoa identificável. Já a *disinformation*, seria o uso consistente e coordenado de informações enganosas ou distorcidas através de diversos expedientes com o intuito de atingir um público definido. Um primeiro grupo utiliza a *disinformation* ao converter cliques em ganho financeiro usando uma variedade de métodos. Outro utiliza com o objetivo de obter vantagem política, avançando suas ideias e atacando adversários políticos. Já um terceiro grupo, mais perigoso, constitui-se de atores estatais que usam a *disinformation* para afetar uma sociedade alvo.¹⁸

Sobre a desinformação patrocinada pelo Estado, visa um público dentro do próprio território do Estado ou no exterior, com fins políticos e estratégicos e o seu impacto tende a ser muito lesivo aos direitos humanos, visto que ao suprimir outras fontes enquanto as “fontes oficiais” promovem suas próprias “falácias”, nega-se o direito das pessoas de buscar e receber informações fidedignas.¹⁹

A análise da desinformação com base em dois vetores – intenção e conhecimento do conteúdo -, também é utilizada.²⁰ De acordo com a intenção, subdivide-se em *misinformation* e *disinformation*, ambas com influências negativas, porém mais graves no segundo caso. Já no que se refere ao conhecimento, divide em informações baseadas em opinião ou baseadas em fatos. As informações falsas baseadas em opinião são criadas para influenciar a opinião ou a decisão dos terceiros. As informações falsas baseadas em fatos são criadas para dificultar a distinção entre informações verdadeiras ou falsas, fazendo com que se acredite na versão falsa, incluindo notícias, rumores e fraudes.

Seguindo uma linha semelhante, o *Facebook*, ao responder o questionário para a formulação do Relatório da ONU sobre desinformação e liberdade de opinião e expressão, revelou que a plataforma adota as seguintes definições para desinformação: a) *Misinformation*: refere-se a conteúdo enganoso (notícias falsas, conteúdo manipulado, etc.); b) *Disinformation*: informações comprovadamente falsas usadas por alguém que sabe que são falsas; c) Operação de influência: esforço coordenado para manipular ou corromper o debate público por um objetivo estratégico. A mídia tenta fazer uma

¹⁶ ANG, Benjamin; ANWAR, Nur Diyanah; JAYAKUMAR, Shashi. *Disinformation & Fake News: Meanings, Present, Future*. p.7

¹⁷ STARTS, Janis. *Disinformation as a Threat to National Security*. p.25

¹⁸ STARTS, Janis. *Disinformation as a Threat to National Security*. p.26

¹⁹ UNITED NATIONS GENERAL ASSEMBLY. *Disinformation and freedom of opinion and expression*. p. 9

²⁰ KUMAR, Srijan; SHAH, Neil. *False Information on Web and Social Media: A Survey*. V.1, n.1 (April 2018). Disponível em <https://arxiv.org/pdf/1804.08559.pdf> Acesso em: 16 mai 2022. p.2

distinção clara entre *disinformation* e *misinformation*, com base no ator, comportamento e conteúdo. Em relação à primeira, foca no sentido de coibir comportamento inautêntico, já na segunda, procura reduzir a sua disseminação.²¹

Interessante, contudo, que o *Facebook*, ao se deparar com esforços coordenados que visam manipular ou corromper o debate público, com objetivo e método estratégico, utiliza a expressão “operações de influência”, conteúdo que ganha especial atenção da referida mídia social.²²

No que tange à liberdade de expressão e opinião, o Comitê de Direitos Humanos da ONU entende que são condições indispensáveis para o pleno desenvolvimento da pessoa, essenciais para qualquer sociedade. A liberdade de pensamento inclui tanto o direito de não ter uma opinião inconscientemente manipulada, além do direito de não revelar pensamentos ou opiniões, bem como não ser penalizado por seus pensamentos.²³

O Relatório da ONU sobre desinformação e liberdade de opinião e expressão aponta que embora a liberdade de opinião seja absoluta, a liberdade de expressão pode sofrer restrições determinadas. A liberdade de opinião possui duas dimensões, uma interna, relacionada à vida privada, à liberdade de pensamento e outra externa, que é justamente a expressão da opinião: a liberdade de expressão. Como a livre circulação de informações é elemento fundamental da liberdade e obrigação do Estado, as únicas possibilidades de restrição à liberdade de expressão são medidas para coibir a violação dos direitos humanos, reputação de terceiros, proteção da segurança nacional, da ordem pública, da saúde ou da moral pública.²⁴

Em 2017 a ONU já havia elaborado declaração conjunta com outras entidades, visando enfrentar o tema, com alguns direcionamentos no sentido de que a limitação ao direito à liberdade de expressão só pode ser permitida com amparo no direito internacional, ou seja, que sejam previstas em lei, que o bem jurídico tutelado seja reconhecido pelo direito internacional, e o tratamento seja movido pelos critérios da necessidade e proporcionalidade. Ainda, as restrições podem ser impostas para proibir a disseminação do ódio, da violência, da discriminação e da hostilidade, nos termos do art.

²¹ FACEBOOK. Facebook submission to UN Special Rapporteur on Freedom of Opinion and Expression for Report on Disinformation, 2021. Disponível em <https://www.ohchr.org/sites/default/files/Documents/Issues/Expression/disinformation/4-Companies/Facebook.pdf>. Acesso em 22 mai. 2022.

²² FACEBOOK. Facebook submission to UN Special Rapporteur on Freedom of Opinion and Expression for Report on Disinformation, 2021.

²³ JONES, Kate. Online Disinformation and Political Discourse: Applying a Human Rights Framework. **International Law Programme**. November, 2019. Disponível em <https://www.chathamhouse.org/sites/default/files/2019-11-05-Online-Disinformation-Human-Rights.pdf> Acesso em 18 mai 2022. p. 34

²⁴ UNITED NATIONS GENERAL ASSEMBLY. Disinformation and freedom of opinion and expression. p. 8

20 (2) do Pacto Internacional sobre Direitos Civis e Políticos.²⁵

Porém, diante das dificuldades de firmar o conceito de desinformação, da lesividade da desinformação sendo difundida em massa, da necessária proteção ao direito de opinião e à liberdade de expressão, a abordagem dos direitos humanos aplicadas à moderação de conteúdo é medida sensata. Pela presença das principais plataformas em vários países, os direitos humanos fornecem uma estrutura jurídica linear e transnacional. Assim, ao invés de discutir a responsabilidade da empresa pelo conteúdo, o ponto de partida é a proteção dos direitos e liberdade individuais.²⁶

Estabelecidas essas premissas, a partir do capítulo seguinte será abordada a lógica da moderação de conteúdo pelas mídias sociais no combate à desinformação.

2. A MODERAÇÃO DE CONTEÚDO PELAS MÍDIAS SOCIAIS NO COMBATE À DESINFORMAÇÃO

Apesar das mídias não serem responsáveis pelo conteúdo, elas possuem um papel fundamental e tomam decisões importantes a respeito da permissibilidade da veiculação, através da moderação de conteúdo, que pode ser compreendida como a prática de determinar quais informações são permitidas e proibidas em suas plataformas (*content gatekeeping*), bem como a forma com que serão classificadas (*content organizing*). Na moderação de conteúdo, as empresas de mídia social contam com uma arquitetura que usa os padrões das comunidades, algoritmos e uma tímida revisão humana.²⁷

A discussão a respeito da responsabilidade pela divulgação do conteúdo iniciou na década anterior ao surgimento das principais mídias sociais hoje no mercado. Evidentemente, o desenvolvimento e massificação do uso das mídias sociais, bem como a concepção do modelo de negócios, foi influenciado por esses casos paradigmáticos. O primeiro caso judicial que analisou o tema com relevância se deu no ano de 1991, quando o Tribunal Distrital dos EUA para o Distrito Sul de Nova York decidiu o caso *Cubby v CompuServe*.²⁸ No caso, venceu o argumento de que a empresa de mídia não possui responsabilidade quando não tem relação direta e prévia com o conteúdo, assim como

²⁵ ORGANIZATION OF AMERICA STATES. Joint declaration of freedom of expression and "Fake News", disinformation and propaganda. 2017.

²⁶ JORGENSEN, Rikke Frank. A Human Rights-Based Approach to Social Media Platforms: Is Social Media Ethical?. 2021. **Berkley Center for Religion, Peace and World Affairs**. Disponível em <https://berkeleycenter.georgetown.edu/responses/a-human-rights-based-approach-to-social-media-platforms>. Acesso em: 16 mar. 2022. p. 2

²⁷ SANDER, Barrie. **Democratic Disruption in the Age of Social Media**: Between Marketized and Structural Conceptions of Human Rights Law. p.163

²⁸ JUSTIA US LAW. *Cubby, Inc. v. CompuServe Inc.*, 776 F. Supp.135 (SDNY 1991). Disponível em <https://law.justia.com/cases/federal/district-courts/FSupp/776/135/2340509/> Acessado em 27 fev. 2022

não possuir relação comercial com os terceiros que veiculavam as informações.²⁹

Por outro lado, no *Stratton Oakmont v Prodigy Services*, de 1995, a Suprema Corte de Nova York entendeu que a empresa que optar por moderar o conteúdo, possui responsabilidade.³⁰ A partir desses precedentes, com o desenvolvimento das grandes plataformas, do avanço da tecnologia e da massiva adesão da população mundial, o tema acerca da responsabilidade das redes tomou corpo, passando pela Seção 230, da *Communications Decency Act*, de 1996, que além de prever a ausência de responsabilidade, contou com a chamada cláusula do bom samaritano, que estatui, em suma, a imunidade das mídias, mesmo para aquelas que optem por moderar conteúdo.³¹

Contudo, por pressão dos usuários e dos governos, várias iniciativas e estudos, privados e governamentais foram desenvolvidos. Da sociedade civil organizada destacam-se as cartas de princípios. De grande importância para a moderação de conteúdo de uma forma geral, destaca-se os Princípios de Santa Clara, que trouxe inicialmente três princípios. O primeiro princípio refere-se ao número, que consiste na necessidade de veiculação do número total de postagens removidas e de contas suspensas, de forma permanente ou temporária, em razão às violações de suas diretrizes de conteúdo. O segundo princípio diz respeito ao aviso, e estatui que as plataformas devem avisar cada usuário cujo conteúdo for removido ou que tiver sua conta suspensa acerca das razões para remoção ou suspensão. O terceiro princípio diz respeito ao recurso, ou seja, as empresas devem fornecer oportunidades significativas para recursos tempestivos acerca de qualquer remoção de conteúdo ou suspensão de conta.³²

Em razão do aumento da importância do papel que essas empresas desempenham na sociedade, fazendo com que a responsabilidade das plataformas seja cada vez maior, organizações e especialistas trabalharam em 2020 e 2021 para desenvolver os Princípios de Santa Clara 2.0., que destacou cinco princípios fundamentais, permanecendo os operacionais os mesmos de 2018: número, aviso e recurso. Os princípios fundamentais de Santa Clara 2.0. preveem a promoção dos direitos humanos e do devido processo legal, de onde se destaca a necessidade de envolver a revisão humana no caso de remoção e suspensão de contas; o estabelecimento de regras claras e compreensíveis; o respeito à competência cultural, que pugna pelo respeito e compreensão do idioma, da cultura e do

²⁹ ESTARQUE, Marina; ACHEGAS, João Victor; BOTTINO, Celina; PERRONE, Christian. Redes Sociais e Moderação de Conteúdo: criando regras para o debate público a partir da esfera privada. **Instituto de Tecnologia e Sociedade do Rio** – **ITS**. Disponível em https://itsrio.org/wp-content/uploads/2021/04/Relatorio_RedemSociaisModeracaoDeConteudo.pdf. Acesso em: 16 mai. 2022. p.17

³⁰ H2O. *Stratton Oakmont, inc. v Prodigy Services Co.* Disponível em <https://h2o.law.harvard.edu/cases/4540> Acesso em: 27 fev. 2022.

³¹ ESTARQUE, Marina; ACHEGAS, João Victor; BOTTINO, Celina; PERRONE, Christian. Redes Sociais e Moderação de Conteúdo: criando regras para o debate público a partir da esfera privada. p.18

³² SANTA CLARA PNCIPLES. Princípios de Santa Clara 1.0., 2018. Disponível em <https://santaclaraprinciples.org/scp1/> Acesso em: 22 mai. 2022.

contexto político e social; o envolvimento do Estado na moderação de conteúdo e, por fim, a integridade e explicação, garantindo que os sistemas de moderação de conteúdo funcionem de maneira confiável.³³

Seguindo essa linha, em 2020, o *Facebook* publicou a *Charting a Way Forward: Online Content Regulation*. Os principais pontos do caminho proposto são: a) Como moderar conteúdo prejudicial e ao mesmo tempo preservar a liberdade de expressão? b) Como as regulamentações devem aumentar a responsabilidade das plataformas da Internet? c) A regulamentação deve exigir que as empresas de Internet cumpram certas metas de desempenho? e) A regulamentação deve definir qual “conteúdo prejudicial” deve ser proibido na internet?³⁴

Ainda do *Facebook*, ao responder o Relatório da ONU sobre desinformação, liberdade de expressão e opinião, revelou que trabalha com três conceitos de desinformação, com já referido: *misinformation*, *disinformation* e operações de influência. Essa última classificação parece inspirar mais cuidados, pois a operação de influência (*influence operation – IO*) se utiliza de contas falsas, páginas e grupos comprometidos com uma de suas faces, a desinformação. Para combater as operações de influência, a plataforma trabalha em dois níveis: 1) Comportamento inautêntico coordenado no contexto doméstico, não governamental; 2) Comportamento inautêntico coordenado por um ator estrangeiro ou governamental.³⁵

Contudo, são os padrões de cada comunidade, fundamentados nos *standards* de direitos humanos, que alimentam a revisão através da inteligência artificial. Os direitos humanos fornecem proteção especial para a opinião e a expressão em assuntos de interesse público como é o caso do direito à informação. Qualquer restrição nesta seara deve obediência aos princípios da legalidade, legitimidade, necessidade e proporcionalidade. O Relatório da ONU sobre liberdade de expressão e opinião traz como exemplo, que a permissão para que as leis proíbam a propagação de falácias relacionadas ao sistema eleitoral e sua integridade devem ser estritas e limitadas a um prazo específico, de forma a não tolher o debate político.³⁶

Na concepção de Jones³⁷, o direito internacional dos direitos humanos prevê o

³³ SANTA CLARA PRINCIPLES. Princípios de Santa Clara 2.0., 2021. Disponível em <https://santaclaraprinciples.org/Acesso em: 22 mai. 2022>

³⁴ FACEBOOK. *Charting a way forward: online content regulations*. Disponível em https://about.fb.com/wp-content/uploads/2020/02/Charting-A-Way-Forward_Online-Content-Regulation-White-Paper-1.pdf Acesso em: 18 mai. 2022.

³⁵ FACEBOOK. Facebook submission to UN Special Rapporteur on Freedom of Opinion and Expression for Report on Disinformation, 2021.

³⁶ UNITED NATIONS GENERAL ASSEMBLY. *Disinformation and freedom of opinion and expression*. p. 8

³⁷ JONES, Kate. *Online Disinformation and Political Discourse: Applying a Human Rights Framework*. p.30

marco legal internacional adequado e disponível para orientar uma atividade empresarial que opera em escala global, impacta na vida das pessoas e possuem poder semelhantes às de um governo local, com alcance ainda maior se considerar o número de afetados.

Dentro da lógica do sistema de moderação de conteúdo, amparado nos *standards* de direitos humanos, urge considerar que o *Twitter* disponibiliza em seu site sua Política Pública para Defesa e Respeito aos direitos das pessoas que utilizam seus serviços.³⁸

O grupo Meta, por sua vez, que reúne o *Facebook*, *Instagram* e *WhatsApp*, possui na sua política, termos de uso e disponibiliza no seu site a sua Política Corporativa de Direitos Humanos³⁹, assim como o *Youtube*, que apresenta na sua página oficial os padrões de Direitos Humanos da *Google*.⁴⁰

Apesar de possuir um parâmetro, as mídias vêm sendo cobradas pela falta de transparência, o que prejudica a proteção dos direitos humanos.⁴¹ De acordo com o Relatório da ONU sobre liberdade de expressão e opinião, as empresas de mídia social contam com suas políticas de conteúdo para lidar com desinformação, contudo, existem dificuldades para encontrar tais regras, isso sem contar com as definições, que geralmente são muito amplas, fazendo com que a aplicação de uma política coordenada de comportamento inautêntico possa ter um efeito adverso sobre a liberdade de expressão. O problema é agravado pelo uso excessivo de filtros automáticos, que possuem dificuldades de captar todas as nuances e a contextualização. Assim, a indeterminação do termo desinformação, os seus elementos políticos, a falta de transparência e a inteligência artificial aumentam o risco de que um conteúdo legítimo seja excluído.⁴²

No que se refere especificamente à desinformação, buscou-se utilizar os parâmetros do *Facebook*⁴³ (2018), maior mídia social do mundo, nos seus Padrões da Comunidade, traz o item “Desinformação” no campo da “Integridade e Autenticidade”. A mídia reconhece de plano que o tratamento da desinformação é diferente dos demais conteúdos, diante das dificuldades de se criar uma lista abrangente de proibições.

A remoção, no *Facebook*, ocorre nos casos de discurso de ódio, quando há risco de agressões físicas, interferência no processo político, acerca da COVID-19, além de uma

³⁸ TWITTER. Defender e respeitar as pessoas que utilizam os nossos serviços. Disponível em <https://help.twitter.com/en/rules-and-policies/defending-and-respecting-our-users-voice> Acesso em: 17 de abr 2022.

³⁹ FACEBOOK. Corporate Human Rights Policy, 2021. Disponível em <https://about.fb.com/wp-content/uploads/2021/03/Facebooks-Corporate-Human-Rights-Policy.pdf>. Acesso em: 18 mai. 2022.

⁴⁰ GOOGLE. Human Rights. 2020. Disponível em <https://about.google/intl/pt-BR/human-rights/> Acesso em: 17 abr. 2022.

⁴¹ IZA, Daniela Gonzales, Human Rights In The Digital Era: Challenges and Opportunities From The United Nations Human Rights System. p. 2

⁴² UNITED NATIONS GENERAL ASSEMBLY. Disinformation and freedom of opinion and expression. p. 15

⁴³ FACEBOOK. Padrões da Comunidade do Facebook, 2018. Disponível em <https://transparency.fb.com/pt-br/policies/community-standards/?source=https%3A%2F%2Fwww.facebook.com%2Fcommunitystandards%2F> Acesso em: 18 mai. 2022.

atenção especial para mídias manipuladas que considera altamente enganosas. Por fim, o *Facebook* revela especial atenção às contas falsas, fraude e comportamento inautêntico e coordenado.

Na sua resposta ao relatório A/HRC/47/25, o *Facebook* revelou ainda a sua estratégia de combate ao conteúdo inautêntico, baseada no ator, comportamento, conteúdo: 1) Atores, envolve a remoção de contas ou organizações pela totalidade de sua atividade na plataforma; 2) Comportamento, se baseia em comportamentos violadores específicos exibidos por atores infratores; 3) Conteúdo, se baseia na aplicação de violações específicas das políticas de conteúdo. E são três as frentes para reduzir a disseminação de notícias falsas e desinformação. Em primeiro lugar, a **remoção**, significando que o *Facebook* remove os conteúdos que violam os padrões da comunidade, o que incluiu contas falsas e envolvidas em comportamento inautêntico, além de desinformação que pode contribuir para o risco de violência iminente ou dano físico, fraude ou interferência eleitoral, discurso de ódio, intimidação e assédio. Por sua vez, a **redução** é dirigida aos conteúdos problemáticos, mas que não violam os padrões da comunidade, sendo que a atuação envolve o rebaixamento no *feed* de notícias. Por fim, a **informação**, que conecta as pessoas com informações precisas evitando a disseminação de informações erradas⁴⁴

Desta forma, após a abordagem da lógica da moderação de conteúdo, importante analisar, no último capítulo, o resultado e as recomendações da ONU contidas no Relatório A/HRC/47/25, sobre liberdade de expressão e opinião no trato da desinformação.

3. O RELATÓRIO DA ONU SOBRE DESINFORMAÇÃO E LIBERDADE DE OPINIÃO E DESINFORMAÇÃO

A Relatora Especial da Organização das Nações Unidas sobre a promoção e proteção do direito à liberdade de opinião e expressão, Irene Khan, após contribuições de 119 organizações da sociedade civil e entidades acadêmicas, 3 organizações internacionais, 3 Estados Membros e 3 empresas, além de consultas às organizações da sociedade civil e suas reuniões com Estados Membros, empresas de mídia e especialistas, editou o A/HRC/47/25 com diversas conclusões e recomendações.

O relatório conclui, inicialmente, que não se deve perder de vista a importância da tecnologia digital para a democracia, de forma que o direito à liberdade de opinião e expressão não fazem parte do problema. Ao mesmo tempo, reconhece a complexidade do fenômeno da desinformação e suas graves consequências para a destruição da confiança

⁴⁴ FACEBOOK. Facebook submission to UN Special Rapporteur on Freedom of Opinion and Expression for Report on Disinformation, 2021.

nas instituições democráticas.⁴⁵

A desinformação, de acordo com o relatório, é um fenômeno complexo e multifacetado com graves consequências para a confiança das pessoas nas instituições democráticas, bem como expõe a facilidade de sua propagação onde a informação pública é deficiente e o jornalismo investigativo é limitado. Enfim, a desinformação é um problema, e são ainda mais problemáticas as repostas que são dadas pelos Estados e pelas empresas, gerando desconfiança dos usuários na integridade da informação⁴⁶. No mesmo sentido, como parte das iniciativas do terceiro princípio da *Paris Call*, o Governo do Canadá realizou seis workshops, com o objetivo de evitar a interferência no processo eleitoral. No primeiro *workshop* trabalhou a ideia de que o compartilhamento de informações de qualidade é fundamental para o combate da interferência da desinformação no processo eleitoral.⁴⁷

Assim, num primeiro plano, o Relatório demonstra sua preocupação com a liberdade de expressão e opinião, apontando serem necessárias respostas coletivas e multidimensionais, sendo os Estados os principais responsáveis por cumprir os direitos humanos. Os Estados, portanto, não devem encorajar ou divulgar alegações que sabem ou deveriam saber que são falsas, nem autorizar o fechamento da Internet como meio de combater a desinformação. A liberdade de expressão não deve ser restringida, com as exceções do disposto nos artigos 19, parágrafo 3, e 20, parágrafo 2, do Pacto Internacional sobre Direitos Civis e Políticos⁴⁸, que devem ser interpretadas de forma estrita e rigorosa⁴⁹.

Quanto às regras das empresas sobre moderação de conteúdo, estas devem ser claras, objetivas e transparentes, de modo que o usuário possa prever a atitude a tomar⁵⁰.

⁴⁵ UNITED NATIONS GENERAL ASSEMBLY. Disinformation and freedom of opinion and expression. p. 17

⁴⁶ UNITED NATIONS GENERAL ASSEMBLY. Disinformation and freedom of opinion and expression. p. 17

⁴⁷ GOVERNMENT OF CANADA. Paris call for trust and security in cyberspace. Disponível em <https://www.canada.ca/en/democratic-institutions/services/paris-call-trust-security-cyberspace.html>. Acesso em: 23 mai. 2022

⁴⁸ Artigo 19. 1. Ninguém poderá ser molestado por suas opiniões. 2. Toda pessoa terá direito à liberdade de expressão; esse direito incluirá a liberdade de procurar, receber e difundir informações e idéias de qualquer natureza, independentemente de considerações de fronteiras, verbalmente ou por escrito, em forma impressa ou artística, ou por qualquer outro meio de sua escolha. 3. O exercício do direito previsto no parágrafo 2 do presente artigo implicará deveres e responsabilidades especiais. Conseqüentemente, poderá estar sujeito a certas restrições, que devem, entretanto, ser expressamente previstas em lei e que se façam necessárias para: a) assegurar o respeito dos direitos e da reputação das demais pessoas; b) proteger a segurança nacional, a ordem, a saúde ou a moral públicas.

Artigo 20. 1. Será proibida por lei qualquer propaganda em favor da guerra. 2. Será proibida por lei qualquer apologia do ódio nacional, racial ou religioso que constitua incitamento à discriminação, à hostilidade ou à violência.

⁴⁹ UNITED NATIONS GENERAL ASSEMBLY. Disinformation and freedom of opinion and expression. p. 18

⁵⁰ JORGENSEN, Rikke Frank. A Human Rights-Based Approach to Social Media Platforms: Is Social Media Ethical? p.2

E mais, o foco da regulação dos meios de comunicação social não deve ser o conteúdo, mas sim a transparência, o direito ao processo legal e o dever de precaução das empresas relativamente aos direitos humanos, garantindo a independência dos órgãos reguladores. Ainda, deve ser priorizada a reorientação do modelo de negócios das mídias, programado para valorizar o sensacionalismo e obter lucro com isso.⁵¹

Neste passo, o *Facebook*, ao submeter suas contribuições ao relatório A/HRC/47/25, reconhecer que a repressão da desinformação pode restringir a liberdade de expressão, alegando que as regras devem ser pautadas na legalidade, necessidade, proporcionalidade para proteger danos a outros direitos, tais como a segurança nacional, a ordem pública e a saúde pública. Ainda, destaca-se das considerações do *Facebook* a revelação de que a plataforma subdivide o combate à autenticidade em três planos: *misinformation*, *disinformation* e operações de influências, essas últimas que oferecem maior gravidade. O Facebook propôs diversas medidas para combater as operações de influência, dentre as quais se destacam a transparência nos anúncios de propaganda política; desenvolvimento de relatórios sobre comportamento inautêntico; sanções econômicas, diplomáticas e/ou criminais; apoio à pesquisa técnica e apoio à alfabetização midiática e digital⁵².

Retomando o relatório, menciona ainda que os Estados devem garantir o direito à informação, aumentando a transparência e divulgando dados oficiais na *Internet* e *off-line*, além de garantir a segurança dos jornalistas. Aponta que a mídia, a informação e a alfabetização digital capacitam as pessoas, preparando-as contra a desinformação, devendo fazer parte do currículo nacional.⁵³ Na declaração conjunta, a ONU já estabelecia a importância da alfabetização midiática e digital.⁵⁴

Acentua novamente o relatório que as mídias devem respeitar os direitos humanos, pois apesar de serem empresas privadas, têm na sua atuação um grande impacto na esfera pública, o que gera uma enorme responsabilidade social. Diante do fato de que são poucas as empresas que dominam o mercado, devem melhorar a moderação de conteúdo e revisar seus modelos de negócios. Devem ainda capacitar os usuários, aumentar a transparência e garantir o devido processo⁵⁵. Nesse ponto, o relatório vai ao encontro das iniciativas concluídas a partir da *Paris Call*, posto que o sexto *workshop* promovido pelo Governo do Canadá, partiu do fundamento de que o cidadão, quando bem

⁵¹ UNITED NATIONS GENERAL ASSEMBLY. Disinformation and freedom of opinion and expression. p. 18

⁵² FACEBOOK. Facebook submission to UN Special Rapporteur on Freedom of Opinion and Expression for Report on Disinformation, 2021.

⁵³ UNITED NATIONS GENERAL ASSEMBLY. Disinformation and freedom of opinion and expression. p. 18

⁵⁴ ORGANIZATION OF AMERICA STATES. Joint declaration of freedom of expression and "Fake News", disinformation and propaganda.

⁵⁵ UNITED NATIONS GENERAL ASSEMBLY. Disinformation and freedom of opinion and expression. p. 18

informado, acaba pensando criticamente, compreendendo os fundamentos da democracia⁵⁶

Sobre o modelo de negócios, as empresas devem garantir que suas atividades comerciais, a práticas de coleta e processamento de dados estejam em conformidade com os padrões internacionais de direitos humanos. Ainda, as empresas devem revisar seus modelos de publicidade, garantindo que não afetem negativamente a diversidade de opiniões e ideias⁵⁷.

Conforme o relatório, as mídias devem adotar políticas de publicidade e conteúdo claras e rigorosamente definidas sobre *Disinformation* e *Misinformation* que estejam de acordo com a lei internacional de direitos humanos. Sobre a sua forma de atuação, as plataformas de mídias sociais devem fornecer informações claras e significativas sobre os parâmetros de seus algoritmos, permitindo que os usuários recebam múltiplas possibilidade de visualizações, aumentando as suas escolhas⁵⁸.

No que tange à transparência, as empresas devem publicar relatórios completos, detalhados e contextualizados, assim como relatórios sobre circunstâncias excepcionais. Os relatórios devem compreender: a) medidas tomadas, b) recursos interpostos; c) número de conteúdos que são compartilhados ou consultados, d) o número de utilizadores alcançados e e) o número de reclamações e pedidos de remoção. Quanto aos recursos, as empresas devem estabelecer mecanismos internos de apelação sobre a moderação, com a criação de mecanismos externos de supervisão⁵⁹. O ponto reafirma os três princípios operacionais de Santa Clara 1.0.: número, aviso e recurso⁶⁰.

As empresas devem ainda priorizar o combate à desinformação de gênero online, dedicar mais recursos para entender melhor os contextos locais que promovem a *Disinformation* e *Misinformation*, com foco nas disparidades de linguagem e conhecimento⁶¹.

Por último, o Relatório realça a importância do sistema de direitos humanos das Nações Unidas, considerando o Conselho de Direitos Humanos fundamental para combater a *disinformation* e *misinformation*, com respeito à liberdade de opinião e expressão, consultando regulamente os Estados, empresas, organizações da sociedade civil e atores, para estabelecer iniciativas de proteção e promoção dos direitos humanos

⁵⁶ GOVERNMENT OF CANADA. Paris call for trust and security in cyberspace.

⁵⁷ UNITED NATIONS GENERAL ASSEMBLY. Disinformation and freedom of opinion and expression. p. 19

⁵⁸ UNITED NATIONS GENERAL ASSEMBLY. Disinformation and freedom of opinion and expression. p. 19

⁵⁹ UNITED NATIONS GENERAL ASSEMBLY. Disinformation and freedom of opinion and expression. p. 19

⁶⁰ SANTA CLARA PRINCIPLES. Princípios de Santa Clara 1.0.

⁶¹ UNITED NATIONS GENERAL ASSEMBLY. Disinformation and freedom of opinion and expression. p. 19

no espaço digital⁶².

CONSIDERAÇÕES FINAIS

O estudo apresentou um panorama dos principais componentes que desafiam o debate acerca do combate à desinformação, a começar pela gravidade do cenário com o avanço da tecnologia e com o modelo de negócio das mídias sociais, pautado na programação de algoritmos para direcionar os usuários às notícias sensacionalistas.

De outro vértice, o estudo demonstrou que o combate à desinformação deve guardar relação com a garantia da liberdade de opinião, direito absoluto e sua subdivisão: a liberdade de expressão. Ambas, protegidas pelo sistema de direitos humanos da Organização das Nações Unidas.

As mídias sociais, muito embora a posição histórica de que num primeiro momento não são responsáveis pelo conteúdo gerado pelos usuários, desenvolveram mecanismos de moderação de conteúdo, a partir dos *standards* de direitos humanos. Contudo, a incerteza da categorização de uma informação como inverídica, somada ao modelo de negócios que lucra com a desinformação e o sensacionalismo, tornam o esforço ainda tímido.

Não há dúvidas da importância dos direitos humanos para a atuação das mídias sociais na moderação de conteúdo, como garantia de um padrão linear e global para atuação em diversos países.

Sobre as contribuições do Relatório A/HRC/47/25 da Relatora Especial da Organização das Nações Unidas sobre a promoção e proteção do direito à liberdade de opinião e expressão, Irene Khan, percebe-se do seu conteúdo e da pesquisa sólida junto a diversos atores, contribuições valiosas para o combate à desinformação através da moderação de conteúdo.

Pode-se destacar do Relatório a preocupação legítima com a liberdade de expressão e de opinião, na medida em que aponta que o combate à desinformação não pode restringir tais principais, ao mesmo tempo em que reconhece que o direito à opinião e expressão são violados a partir do grande fluxo de notícias falsas e coordenadas.

Nesta seara, ainda, o Relatório reafirma que a única possibilidade de restrição à liberdade de expressão são as medidas para coibir a violação dos direitos humanos, reputação de terceiros, proteção da segurança nacional, da ordem pública, da saúde ou da moral pública, com base no Pacto Internacional sobre Direitos Civis e Políticos.

Outra importante contribuição é a respeito do foco em que deve se dar a regulação.

⁶² UNITED NATIONS GENERAL ASSEMBLY. Disinformation and freedom of opinion and expression. p. 20

Ao invés do Estado centrar esforços na regulação pura e simples do conteúdo das postagens, deve se preocupar com a transparência, com o direito ao processo legal e com o dever de precaução das empresas com os direitos humanos.

Importante ainda a revisão no modelo de negócios das empresas, de forma que os algoritmos e o sistemas de classificação, ao invés de amplificar a desinformação, apresentem uma múltipla possibilidade de visualizações, aumentando as escolhas dos usuários quanto à informação.

Por fim, o relatório menciona que o Conselho de Direitos Humanos da ONU é peça importante para o combater a desinformação, ao tempo em que contribuirá para garantir a liberdade de opinião e expressão, com a interação periódica entre os Estados, as mídias e os demais atores.

Como se verificou, o Relatório A/HRC/47/25 da Relatora Especial da Organização das Nações Unidas sobre a promoção e proteção do direito à liberdade de opinião e expressão, Irene Khan, trouxe em larga medida, importantes contribuições para o tema, muitas delas expressando o posicionamento da sociedade civil organizada, especialmente nos Princípios de Santa Clara e na Carta de Paris. Contudo, o esforço para a melhora do cenário é permanente e exige o esforço integrado de diversos atores: Estados, sociedade civil, academia e as mídias sociais.

REFERÊNCIAS DAS FONTES CITADAS

ANG, Benjamin; ANWAR, Nur Diyanah; JAYAKUMAR, Shashi. Disinformation & Fake News: Meanings, Present, Future, 2018. Em **Disinformation and Fake News. 3-20. Singapore: Palgrave Macmillan**. DOI: <https://doi.org/10.1007/978-981-15-5876-4>

AIETA, Vania Siciliano. O Impacto Eleitoral Resultante da Manipulação das Fake News no Universo das Redes Sociais: a Construção da Desinformação. **Revista Interdisciplinar de Direito**. v. 18, n. 1, pp 213-233, jan./jun. 2020.

BARBOSA, Mafalda Miranda. Fake News e Fact-checkers: uma perspectiva jurídico-civilista. **Revista de Direito e Responsabilidade**. Ano 3 – 2021. p. 733-766

ESTARQUE, Marina; ACHEGAS, João Victor; BOTTINO, Celina; PERRONE, Christian. Redes Sociais e Moderação de Conteúdo: criando regras para o debate público a partir da esfera privada. **Instituto de Tecnologia e Sociedade do Rio – ITS**. Disponível em https://itsrio.org/wp-content/uploads/2021/04/Relatorio_RedetesSociaisModeracaoDeConteudo.pdf. Acesso em: 16 mai. 2022.

FACEBOOK. Charting a way forward: online content regulations. Disponível em <https://about.fb.com/wp-content/uploads/2020/02/Charting-A-Way-Forward-Online-Content-Regulation-White-Paper-1.pdf> Acesso em: 18 mai. 2022.

FACEBOOK. Corporate Human Rights Policy, 2021. Disponível em <https://about.fb.com/wp-content/uploads/2021/03/Facebooks-Corporate-Human-Rights-Policy.pdf>. Acesso em: 18 mai. 2022.

FACEBOOK. Facebook submission to UN Special Rapporteur on Freedom of Opinion and Expression for Report on Disinformation, 2021. Disponível em <https://www.ohchr.org/sites/default/files/Documents/Issues/Expression/disinformation/4-Companies/Facebook.pdf>. Acesso em 22 mai. 2022.

FACEBOOK. Padrões da Comunidade do Facebook, 2018. Disponível em <https://transparency.fb.com/pt-br/policies/community-standards/?source=https%3A%2F%2Fwww.facebook.com%2Fcommunitystandards%2F> Acesso em: 18 mai. 2022.

GOOGLE. Human Rights. 2020. Disponível em <https://about.google/intl/pt-BR/human-rights/> Acesso em: 17 abr. 2022.

GOVERNMENT OF CANADA. Paris call for trust and security in cyberspace. Disponível em <https://www.canada.ca/en/democratic-institutions/services/paris-call-trust-security-cyberspace.html>. Acesso em: 23 mai. 2022

H2O. Stratton Oakmont, inc. v Prodigy Services Co. Disponível em <https://h2o.law.harvard.edu/cases/4540> Acesso em: 27 fev. 2022.

IZA, Daniela Gonzales. Human Rights In The Digital Era: Challenges and Opportunities From The United Nations Human Rights System. **Jurnal Pengajian Umum Asia Tenggara** 22 (2021): 1 – 13

JONES, Kate. Online Disinformation and Political Discourse: Applying a Human Rights Framework. **International Law Programme**, November 2019. Disponível em <https://www.chathamhouse.org/sites/default/files/2019-11-05-Online-Disinformation-Human-Rights.pdf> Acesso em 18 mai 2022.

JORGENSEN, Rikke Frank. A Human Rights-Based Approach to Social Media Platforms: Is Social Media Ethical?. 2021. **Berkley Center for Religion, Peace and Word Affairs**. Disponível em <https://berkeleycenter.georgetown.edu/responses/a-human-rights-based-approach-to-social-media-platforms>. Acesso em: 16 mar. 2022.

JUSTIA US LAW. Cubby, Inc. v. CompuServe Inc., 776 F. Supp. 135 (SDNY 1991). Disponível em <https://law.justia.com/cases/federal/district-courts/FSupp/776/135/2340509/> Acessado em 27 fev. 2022

KUMAR, Srijan; SHAH, Neil. False Information on Web and Social Media: A Survey. 1, 1 (April 2018). Disponível em <https://arxiv.org/pdf/1804.08559.pdf> Acesso em: 16 mai 2022.

MANYANA, Candice Chirwa and Zimkhitha. The Rise of Fake News: Surveying the Effects of Social Media on Informed Democracy. **The Thinker**. Volume 88, 2021. Disponível em https://journals.uj.ac.za/index.php/The_Thinker/article/view/604/378. Acesso em: 18 mai. 2022.

ORGANIZATION OF AMERICA STATES. Joint declaration of freedom of expression and “Fake News”, disinformation and propaganda. 2017. Disponível em <https://www.osce.org/files/f/documents/6/8/302796.pdf>. Acessado em 15 de mai. 2022.

SANDER, Barrie. Democratic Disruption in the Age of Social Media: Between Marketized and Structural Conceptions of Human Rights Law. **The European Journal of International Law**. Vol. 32 no. 1, 2021.

SANTA CLARA PRINCIPLES. Princípios de Santa Clara 1.0., 2018. Disponível em <https://santaclaraprinciples.org/scp1/> Acesso em: 22 mai. 2022.

SANTA CLARA PRINCIPLES. Princípios de Santa Clara 2.0., 2021. Disponível em <https://santaclaraprinciples.org/> Acesso em: 22 mai. 2022.

STARTS, Janis. Disinformation as a Threat to National Security. 2018. Em **Disinformation and Fake News**. 23-32. Singapore: Palgrave Macmillan. DOI: <https://doi.org/10.1007/978-981-15-5876-4>

TWITTER. Defender e respeitar as pessoas que utilizam os nossos serviços. Disponível em <https://help.twitter.com/en/rules-and-policies/defending-and-respecting-our-users-voice> Acesso em: 17 de abr 2022.

UNITED NATIONS GENERAL ASSEMBLY. Disinformation and freedom of opinion and expression. Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Irene Khan A/HRC/47/25. Disponível em <https://www.ohchr.org/en/calls-for-input/reports/2021/report-disinformation>

VÉRITER, Sophie. European Democracy and Counter-Disinformation: Toward a New Paradigm?. **Carnegie Europe**. 2021. Disponível em: <https://carnegieeurope.eu/2021/12/14/european-democracy-and-counter-disinformation-toward-new-paradigm-pub-85931>>. Acesso em: 24 fev. 2022.

COMO CITAR:

BORGES, Gustavo Silveira; SOUZA, Fabio Jeremias de. Desinformação, liberdade de opinião e expressão e moderação de conteúdo nas mídias sociais. **Revista Direito e Política**. Programa de Pós-Graduação Stricto Sensu em Ciência Jurídica da UNIVALI, vº 19, nº3, 3º quadrimestre de 2024. Disponível em: <https://periodicos.univali.br/index.php/rdp> - ISSN 1980-7791. DOI: <https://doi.org/10.14210/rdp.v19n3.p353-373>

INFORMAÇÕES DOS AUTORES:

Gustavo Silveira Borges

Pós-Doutor em Direito Civil, com bolsa de pesquisa PNPd/CAPES, na Universidade do Vale do Rio dos Sinos - UNISINOS (2014). Doutor em Direito Civil na Universidade Federal do Rio Grande do Sul - UFRGS (2013). Mestre em Ciências Criminais pela Pontifícia Universidade Católica do Rio Grande do Sul - PUCRS (2007). Especialista em Ciências Penais pela PUCRS (2005). Pós-graduação pela Escola da Magistratura do Rio Grande do Sul - AJURIS (2003). Graduado em Ciências Jurídicas e Sociais pela PUCRS (2002). Atualmente, é Professor da Graduação e Pós-Graduação da Faculdade de Direito da Universidade do Extremo Sul Catarinense e da Pós-graduação do Centro Universitário Ritter dos Reis - Uniritter. Parecerista em diversas revistas. Participante em diversos projetos de pesquisa vinculados à PUCRS e UFRGS. Pesquisador do Núcleo de Pesquisas em Direitos Humanos e Cidadania - NUPEC, na UNESCO. Autor de diversos trabalhos científicos desenvolvidos nas áreas do Direito e da Medicina. Tem experiência no Direito, com ênfase em Direito Civil e Processual Civil, com atuação acadêmica nas áreas: Direito Civil, Direito Processual Civil, Direito do Consumidor, Responsabilidade Civil, Contratos e Direito Médico. Advogado e parecerista em Direito Médico.

Fabio Jeremias de Souza

Mestrando em Direito pela Universidade do Extremo Sul de Santa Catarina. Graduação em Direito pela Universidade do Sul de Santa Catarina (2000). Especialização em Pós Graduação em Direito Empresarial pela Universidade do Extremo Sul Catarinense, Brasil(2003) Atualmente é advogado - Escritório de Advocacia. Vice-Presidente da Subseção de Criciúma/SC da OAB. (Gestão 13/15) Presidente da Subseção de Criciúma/SC da OAB. (Gestão 16/18) Conselheiro Federal da OAB. (Gestão 19/22) Membro fundador da Academia Catarinense de Direito Eleitoral

Recebido em: 23/07/2022
Aprovado em: 14/02/2024

Received: 23/07/2022
Approved: 14/02/2024