

TRANSPARÊNCIA E CONTROLE DE CONTEÚDO EM REDES SOCIAIS E O DIREITO FUNDAMENTAL À LIBERDADE DE EXPRESSÃO: ANÁLISE DA ATUAÇÃO DO *OVERSIGHT BOARD* NO *FACEBOOK* (META)¹

TRANSPARENCY AND CONTENT CONTROL IN SOCIAL MEDIA AND THE FUNDAMENTAL RIGHT TO FREEDOM OF EXPRESSION: ANALYSIS OF THE PERFORMANCE OF THE OVERSIGHT BOARD REGARDING FACEBOOK (META)

Artigo recebido em 23/04/2022

Artigo aceito em 10/05/2022

Artigo publicado em 29/01/2023

Marco Aurélio Marrafon

Professor da Faculdade de Direito da Universidade do Estado do Rio de Janeiro – UERJ (graduação, mestrado e doutorado). Doutor e Mestre em Direito do Estado pela Universidade Federal do Paraná - UFPR, com estudos doutorais na Università degli Studi Roma Tre - Italia. Coordenador do Laboratório de Direito e Inteligência Artificial – LabDIA da FD/UERJ. Presidente (2012-2018) e Membro da Academia Brasileira de Direito Constitucional - ABDConst. Membro da Comissão Nacional de Estudos Constitucionais do CFOAB. Advogado. *ORCID*: <https://orcid.org/0000-0002-6891-6221>. *E-mail*: marco_marrafon@yahoo.com.br.

Marina Giovanetti Lili Lucena

Doutoranda em Direito Civil na Universidade do Estado do Rio de Janeiro – UERJ. Mestre em Direito e Inovação pela Universidade Federal de Juiz de Fora - UFJF. Graduada em Direito pela UFJF, com período de intercâmbio acadêmico na Università degli Studi di Camerino – Italia. *ORCID*: <https://orcid.org/0000-0002-7092-0190>. *E-mail*: marinagiovanetti@gmail.com.

RESUMO: O presente estudo tem por objetivo principal examinar o Oversight Board, sua composição, atuação e efetividade. Tal órgão foi criado pela empresa Meta com o intuito de realizar análise posterior de alguns casos de moderação de conteúdo de plataformas como Facebook e Instagram. Parte-se do pressuposto de que, em contexto de grande proliferação das redes sociais e dos conteúdos nelas compartilhados, torna-se mais complexo definir o que pode permanecer e o que excede o direito à liberdade de expressão. A questão torna-se ainda mais complexa considerando o grande número de usuários ao redor do mundo nessas plataformas. Além disso, muitas vezes essas redes sociais não são transparentes sobre seu funcionamento, tornando-se uma aplicação concreta da chamada Black Box. Nesse sentido, serão analisados conceitos como a liberdade de expressão, moderação de conteúdo e transparência, além de verificar como o órgão independente criado pela Meta pode auxiliar na compatibilização de tais conceitos nos casos concretos.

¹ Artigo produzido no âmbito do Grupo de Pesquisas Institucional “Novas Tecnologias, Inteligência Artificial, Direito e Democracia”, vinculado ao Programa de Pós-graduação em Direito da Universidade do Estado do Rio de Janeiro – UERJ, sob Coordenação do Prof. Dr. Marco Aurélio Marrafon. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) – Código de Financiamento 001.

PALAVRAS-CHAVE: Liberdade de expressão. Black Box. Meta. Moderação de conteúdo. Oversight Board. Transparência.

ABSTRACT: The main objective of this study is to examine the Oversight Board, its composition, performance, and effectiveness. Such a body was created by the company Meta in order to perform further analysis of some cases of content moderation on platforms like Facebook and Instagram. It is assumed that, in the context of a great proliferation of social media platforms and the content shared on them, it becomes more complex to define what can remain on them and what exceeds the right to freedom of expression. The issue becomes even more complex considering the large number of users around the world on these platforms. Moreover, these social networks are often not transparent about their functioning, becoming a concrete application of the so-called Black Box. In this sense, we will analyze concepts such as freedom of expression, content moderation, and transparency, and verify how the independent body created by Meta can help in creating compatibility between such concepts in concrete cases.

KEYWORDS: Freedom of expression. Black Box. Meta. Content moderation. Oversight Board. Transparency.

SUMÁRIO: Introdução. 1 Facebook, Meta e a Black Box Society. 2 Liberdade de expressão e moderação de conteúdo. 3 O Oversight Board. 4 Breve inventário das decisões do Oversight Board. 5 Avaliação Preliminar. Conclusão. Referências bibliográficas.

TABLE OF CONTENTS: Introduction. 1 Facebook, Meta and the Black Box Society. 2 Freedom of expression and content moderation. 3 The Oversight Board. 4 Brief assessment of the Oversight Board's decisions. 5 Preliminary Assessment. Final Considerations. Bibliographical references.

1 INTRODUÇÃO

Atualmente, a sociedade e o Direito se organizam de modo a tentar compreender as transformações ocorridas com as novas tecnologias e assim, realizar as mudanças e adaptações necessárias. Nas últimas décadas, o desenvolvimento das tecnologias

comunicacionais trouxe novos dilemas e questionamentos, sendo essencial para o presente trabalho a criação e fortalecimento das chamadas redes sociais.

As plataformas de redes sociais permitem a comunicação entre indivíduos distantes fisicamente, o que ocorre de maneira barata, instantânea e em escala inimaginável há algumas décadas. Nelas, as pessoas podem se comunicar tanto com amigos quanto com desconhecidos, sendo possível a criação e compartilhamento de conteúdo em diversos formatos – fotos, vídeos, áudios e textos. Nesse contexto, intensificaram-se as discussões sobre a necessidade de algum tipo de controle dos conteúdos compartilhados nessas plataformas sociais. A questão torna-se ainda mais relevante quando o material produzido e compartilhado é considerado lesivo, a exemplo dos casos de desinformação e ofensa a alguém.

Notadamente em tempos de pandemia ou mesmo no período eleitoral, conteúdos distorcidos, notícias falsas e teorias da conspiração tem o potencial de causar catástrofes humanitárias ou mesmo graves danos à democracia e aos pilares basilares do Estado de Direito, dando ensejo à emergência de discursos autoritários e populistas².

Nesse contexto, é preciso avançar na compreensão teórica das possibilidades de controle e moderação de conteúdo em redes sociais, de modo a evitar que esse controle se torne uma espécie de censura prévia vedada constitucionalmente e se garanta a liberdade de expressão protegida pela Constituição (art. 5º, IX, CF/88).

O debate é incipiente e necessita de referências teóricas e práticas para que a atividade moderadora se realize de maneira equilibrada entre o direito fundamental à liberdade de expressão e o direito à informação verdadeira e de qualidade, coibindo maiores danos às pessoas, à sociedade e à democracia constitucional.

Nesse diapasão, o presente artigo busca compreender a atuação de controle do *Oversight Board* (ou Comitê de Supervisão) da plataforma *Facebook* da empresa Meta e de modo a apresentar um breve inventário de suas decisões, avaliando-as criticamente, a fim de colher subsídios que acerca de sua transparência e *accountability* e, quiçá possam contribuir para a elaboração da regulamentação legislativa ordinária e constitucional do tema, à luz da experiência concreta da regulação privada.

Com efeito, é comum que a regularização desses conteúdos ocorra por regulações específicas e privadas de cada plataforma. Essas regras são determinadas, essencialmente, pelos chamados Termos de Uso. A análise do presente artigo irá se centrar na atuação da

² Sobre o tema, conferir: ROBL FILHO, Ilton Norberto; MARRAFON, Marco Aurélio; PANSIERI, Flávio. Constitucionalismo como salvaguarda do Estado de Direito: crítica ao (ciber) populismo autoritário e a necessária reengenharia constitucional. Revista do Instituto de Hermenêutica Jurídica: RIHJ., v. 18, 2020.

empresa Meta, composta, dentre outras, pela rede social *Facebook*. A empresa possui grande poder econômico e social, com grande quantidade de conteúdos e usuários presentes diariamente na plataforma. No entanto, a criação recente do *Oversight Board*, se pretende como uma exceção a dinâmica da empresa de atuar como agente econômico que cria as próprias regras e as aplica em seus julgamentos.

Nesse sentido, inicialmente será analisada a plataforma *Facebook*, a criação da empresa Meta e a forma como sua atuação pode ser entendida como integrante da lógica de *Black Box*, ou seja, de ausência de transparência para o público em geral. Posteriormente, será estudada a ideia de moderação de conteúdo, normalmente realizada pelas redes sociais, e sua relação com o direito fundamental à liberdade de expressão. Os limites e conflitos entre a liberdade de expressão e outros direitos é temática essencial na atualidade, principalmente no mundo *online*.

Será ainda realizado exame sobre o Comitê de Supervisão, sua criação e composição, e a forma como realiza suas análises. Serão investigadas as decisões disponibilizadas por esse ente, de modo a compreender suas dinâmicas de decisão. Nesse sentido, pretende-se responder a questões sobre a publicidade e o acesso dessas decisões, quais informações são disponíveis ao público e como há essa disponibilização.

Assim, será possível compreender se a atuação do *Oversight Board* pode ser visualizada como uma análise independente e cujo modelo pode ser replicado em ambiente público mais amplo, guardadas as devidas conformações. Além disso, pretende-se entender se sua atuação é mecanismo apto para diminuir a *Black Box* que hoje vigora nas redes sociais e, como tal, se auxilia na transparência das redes sociais da empresa Meta.

A pesquisa documental terá como fontes principais normas jurídicas, termos de uso de plataformas sociais e decisões do *Oversight Board*. Esses serão analisados como fontes primárias, pela análise de documentos originais, sem intermediação (GUSTIN; DIAS, 2013, p. 75). Tais documentos serão acessados essencialmente por meio virtual³. As fontes secundárias, por sua vez, serão artigos e livros acadêmicos (REGINATO, 2017, p. 198), resultado das análises de outros pesquisadores (GUSTIN; DIAS, 2013, p. 75). Como revisão bibliográfica, para formulação da fundamentação teórica, serão utilizadas referências sobre o desenvolvimento da internet, as redes sociais, a regulação digital, o direito à liberdade de expressão e a ausência de transparência das plataformas sociais.

³ Sua utilização é válida, desde que satisfeitas as checagens de validade e idoneidade da fonte (REGINATO, Andréa Depieri de A. Uma introdução à pesquisa documental. In: MACHADO, Maíra Rocha (Org.). *Pesquisar empiricamente o direito*. São Paulo: Rede de Estudos Empíricos em Direito, 2017 p. 216).

2 FACEBOOK, META E A BLACK BOX SOCIETY

A sociedade inegavelmente passou por transformações sociais, econômicas e tecnológicas nas últimas décadas. Uma das questões que se transformou de forma notória foi a comunicação social. Dessa maneira, o aprimoramento tecnológico possibilitou que novas formas e modalidades de interação entre indivíduos se desenvolvessem, permitindo a comunicação fácil, rápida e barata, mesmo em casos de distância física entre os indivíduos.

Como forma de interação social, tornou-se comum a criação e desenvolvimento das chamadas redes sociais, que permitem a criação e compartilhamento de conteúdo, seja aquele próprio ou de autoria de terceiros. Elas, com sua grande influência na atualidade, concretizam uma revolução da mídia. Com os avanços tecnológicos, houve ampliação dos meios tradicionais de comunicação, com a criação de novos espaços de circulação de informações (SCHREIBER, 2013, p. 11-12).

Para fazer parte dessas redes, cada indivíduo deve criar o seu perfil individual. Tal processo faz parte da construção da identidade virtual ou corpo eletrônico (RODOTÀ, 2012, p. 159), ou seja, a visão pública criada a partir de um perfil, pela junção de informações pessoais exigidas, a exemplo de foto e nome. A exposição de conteúdos diversos faz parte do processo de publicização da vida privada, com a intimidade individual sendo cada vez mais aberta ao público (BRANCO, 2017, p. 32). A chamada internet 2.0 é resultado desse desenvolvimento, com a característica marcante de utilização de várias redes sociais (BRANCO, 2017, p. 319).

Assim, redes sociais são plataformas que permitem a criação e circulação de conteúdo próprio ou alheio, que pode ser de áudio, audiovisual ou textual (KELLER, 2019, p. 258-259). Nesses espaços são compartilhadas questões sobre política, eleição, memes⁴, músicas, fotos e danças, dentre outras inúmeras possibilidades temáticas.

No presente trabalho será enfocada a atuação das redes sociais pertencentes à empresa Meta. Em outubro de 2021 Mark Zuckerberg apresentou a Meta, empresa que reúne os aplicativos e tecnologias sob essa nova marca⁵. São aplicativos reunidos o *Facebook*, *Messenger*, *Instagram*

⁴ Memes são conteúdos humorísticos, geralmente vinculado a imagens (de pessoas, ambientes e obras de arte, dentre outras), que são compartilhadas nas redes.

⁵ A apresentação da empresa de tecnologia digital encontra-se disponível em: <https://about.fb.com/br/news/2021/10/apresentando-meta-uma-empresa-de-tecnologia-social/>. Acesso em: 09 jan. 2022.

e *WhatsApp*. O nome Meta decorre da ideia de metaverso, tecnologia que promete se desenvolver nos próximos anos e revolucionar a forma de comunicação entre os indivíduos⁶.

Assim, neste trabalho será referido à Meta como a empresa de tecnologia que se utiliza do *Oversight Board*. O *Facebook* será mencionado como a rede social utilizada pelos indivíduos para se conectar com amigos. Nesse sentido se posiciona o próprio Comitê de Supervisão que, a partir do caso 2021-012-FB-UA, julgado em 09 de dezembro de 2021, passa a se referir nas decisões à Meta como empresa, e ao *Facebook* como aplicativo específico, com políticas associadas. Como será visto, o *Oversight Board* já se manifestou sobre decisões do *Instagram* e do *Facebook*.

Salienta-se que a mudança de nome da empresa parece se desenvolver em contexto de certa crise do *Facebook*. Inicialmente, houve o escândalo *Cambridge Analytica*⁷, com suspeita de envolvimento do *Facebook* no uso eleitoral de dados pessoais. Posteriormente, a invasão do Capitólio no início de 2021 mostrou que condutas planejadas digitalmente podem afetar as democracias de forma séria. Havia suspeita de que essas ações foram planejadas por redes sociais, incluindo o *Facebook* (ALBUQUERQUE, 2021).

Ainda em 2021, a ex-funcionária do *Facebook* Frances Haugen fez denúncias sérias sobre a atuação da empresa, incluindo a utilização de algoritmos que incentivam a discórdia e a utilização de mecanismos para viciar os usuários. Em resumo, a denúncia trazia problemas sociais causados pelo *Facebook* na busca de maiores lucros. Posteriormente, Haugen compareceu ao Senado dos Estados Unidos para exigir mudanças (SEISDEDOS, 2021). Já em 2022, a Meta perdeu mais de 230 bilhões de dólares em valor de mercado, no que se concretizou como a pior sessão da sua história, incluindo a diminuição do número de usuários ativos⁸.

Assim, após as várias denúncias e explicações exigidas do *Facebook* e da Meta, como visto em alguns episódios supracitados, a empresa parece buscar maior transparência e *accountability* das suas ações, o que tenta ser concretizado, ao menos em parte, pelo *Oversight Board*.

⁶ Disponível em: <https://about.facebook.com/br/meta/>. Acesso em: 09 jan. 2022.

⁷ O caso *Cambridge Analytica* denunciou o uso eleitoral de dados pessoais, principalmente pelo *Facebook*. Os dados de cerca de 50 milhões de pessoas foram obtidos pela realização de testes de personalidade nesta rede social. Posteriormente, foram utilizados para construir um perfil dos eleitores e realizar *microtargeting* comportamental. Suspeita-se que os resultados tenham sido utilizados na campanha de Donald Trump para a presidência dos Estados Unidos da América em 2016. Sobre isso: SILVA, Gabriel Souza de. *O caso Cambridge Analytica: distorções na comunicação política e retração da esfera pública em arenas discursivas digitais*. Rio de Janeiro: Lumen Juris, 2021.

⁸ A informação pode ser acessada em: <https://www.infomoney.com.br/mercados/acoes-do-facebook-fbok34- agora-meta-desabam-22-apos-balanco-com-numeros-abaixo-do-esperado/>. Acesso em: 09 fev. 2022.

Nesse sentido, conceito essencial para o seguinte trabalho é o desenvolvido por Frank Pasquale no livro *The Black Box Society* (2016). O termo *Black Box* é colocado pelo autor como uma metáfora, dual, ao dispositivo de rastreamento (em aviões, por exemplo) e o funcionamento misterioso do sistema (2016, p. 3). A ideia do termo *Society*, por sua vez, é para reafirmar que a distinção fixa e definitiva entre o mercado e o Estado está gradualmente desaparecendo (2016, p. 10).

Pasquale também questiona um paradoxo na sociedade atual: enquanto a privacidade individual é cada vez mais violada, com todas as ações humanas sendo armazenadas e catalogadas, há poderosas empresas, inclusive na área da internet, que escondem suas ações e mecanismos de escolhas. Um dos seus principais objetivos é tratar das empresas do Vale do Silício (2016, p. 6). Dessa maneira, é possível incluir a Meta.

O autor questiona que alguns ambientes, que são espaços da nossa vida comum e em sociedade, deveriam ser mais abertos e transparentes. No entanto, argumenta que tais espaços muitas vezes são determinados pela lógica de segredo. O autor menciona que o funcionamento desses sistemas é misterioso e que a autoridade dos algoritmos para tomada de decisões está aumentando. Desse modo, as decisões que antes eram humanas tornam-se, muitas vezes, automatizadas (2016, p. 8), ou seja, realizadas pelos algoritmos. Porém, a sociedade não sabe (e não tem como saber) se o algoritmo é justo, porque seu mecanismo de funcionamento costuma ser secreto (PASQUALE, 2016, p. 9).

Ponto interessante salientado pelo autor é que a transparência não é suficiente, sendo necessário que vá além, alcançando o ponto da inteligibilidade. Nesse sentido, a abertura não pode ser um gesto vazio, em que a complexidade das informações disponibilizadas torne as transações inexplicáveis aos indivíduos (PASQUALE, 2016, p. 8 e 16).

O presente trabalho analisa o processo de retirada ou manutenção de conteúdo em redes sociais da empresa Meta. A escolha, por vezes algorítmica, de inserir ou não determinado conteúdo como violador dos Padrões da Comunidade pode trazer prejuízos ao indivíduo. Tais prejuízos ocorrem tanto para a sua liberdade de expressão, quanto podem causar lesões financeiras, já que muitas pessoas desenvolvem sua vida profissional na internet e nas redes sociais.

Ao trazer contribuições sobre como tornar a sociedade mais transparente, Pasquale argumenta que o foco deveria ocorrer na maneira como as empresas tomam as suas decisões (2016, p. 141). Uma possibilidade seria humanizar o processo, com revisões internas e direitos de

recurso externo (PASQUALE, 2016, p. 198). Por isso, analisa-se se a atuação do *Oversight Board* nas decisões sobre conteúdo realizadas pela Meta concretiza essa maior transparência.

3 LIBERDADE DE EXPRESSÃO E MODERAÇÃO DE CONTEÚDO

A atividade de análise dos limites da liberdade de expressão no mundo *online* e moderação de conteúdo pelas plataformas é temática com relevância crescente na sociedade, inclusive juridicamente. Como visto, as redes sociais e plataformas de conteúdo permitem compartilhamento de variados assuntos, em vários formatos e, por vezes, esse compartilhamento pode ser lesivo aos demais.

No Brasil, a liberdade de expressão é positivada na Constituição Brasileira de 1988 (CF/88), que buscou garantir esse direito fundamental após período ditatorial. Sua previsão ocorre no artigo 5º, IX⁹ e 220¹⁰. Internacionalmente, a liberdade de expressão também é direito consagrado, a exemplo da sua previsão no artigo 19 da Declaração Universal dos Direitos Humanos¹¹.

No entanto, salienta-se que a liberdade de expressão não significa liberdade absoluta de manifestação. Como qualquer direito, as escolhas que são realizadas pelo indivíduo devem se coadunar com aquelas feitas pelos demais membros da sociedade, ou seja, “a liberdade de cada um deve conviver com a liberdade de todos” (BINENBOJM, 2020, p. 14). Não se pode esquecer também o papel coletivo e democrático desse direito e, assim, seu exercício não pode produzir lesões a outros bem jurídicos também protegidos pelo ordenamento jurídico nacional, sendo necessária a imposição de limites (MORAES; KONDER, 2012, p. 3).

Nesse sentido se manifesta o próprio *Oversight Board*, segundo o qual “a liberdade de expressão é fundamental, mas há momentos em que o discurso pode violar a autenticidade, segurança, privacidade e dignidade. Algumas expressões podem comprometer a capacidade de outras pessoas se expressarem livremente”¹².

⁹ Art. 5º, IX, CF/88: IX - é livre a expressão da atividade intelectual, artística, científica e de comunicação, independentemente de censura ou licença.

¹⁰ Art. 220, CF/88: Art. 220. A manifestação do pensamento, a criação, a expressão e a informação, sob qualquer forma, processo ou veículo não sofrerão qualquer restrição, observado o disposto nesta Constituição.

¹¹ *In verbis*: “Todo ser humano tem direito à liberdade de opinião e expressão; esse direito inclui a liberdade de, sem interferência, ter opiniões e de procurar, receber e transmitir informações e ideias por quaisquer meios e independentemente de fronteiras”.

¹² Disponível em: <https://oversightboard.com/governance/>. Acesso em: 10 fev. 2022.

Desse modo, como todos os outros direitos, a liberdade de expressão é determinada por outras normas, jurídicas ou não (ANDRADE, 2020, p. 115). A título exemplificativo, cita-se a questão sobre o conteúdo produzido por terceiro, de grande relevância no mundo *online*. O Brasil regulou tal questão por via legislativa no Marco Civil da Internet (MCI) que, nos artigos 18 a 21, estabelece a responsabilidade civil dos provedores de conexão à internet e de aplicações de internet¹³.

Em resumo, o direito à liberdade de expressão é fundamental, mas abusos e lesões a outros direitos não podem ser tolerados. Atualmente, resta incontestado que esse direito não é ilimitado, ainda que, por vezes, sejam complexas as determinações sobre as suas limitações, essencialmente nas redes sociais. No caso da moderação de conteúdo, essas normas costumam ser privadas, ou seja, estabelecidas pela própria empresa.

A atividade de moderação de conteúdo pelas plataformas pode envolver atividades diversas. Em um primeiro ponto, é comum que essas plataformas estabeleçam critérios para a moderação, o que ocorre principalmente nos Termos de Uso ou os Padrões da Comunidade. No presente trabalho serão mencionados, de maneira geral, as normas definidas por duas redes sociais que pertencem à Meta e que já foram alvo de decisões do *Oversight Board*, quais sejam, *Instagram* e *Facebook*.

Os Termos de Uso do *Instagram*¹⁴ são disponíveis *online* e foram atualizados em 04 de janeiro de 2022, em razão da mudança do nome da empresa para Meta. É informada a utilização de ferramentas para combater abusos e violações e promover um ambiente seguro, inclusivo e positivo. Na parte sobre remoção de conteúdo, há informações sobre a remoção de conteúdos que violem os termos de Uso ou as políticas, incluindo as Diretrizes da Comunidade do *Instagram*, além dos casos de autorização e obrigatoriedade derivadas da lei.

¹³ Para entender a responsabilidade civil desses provedores, é essencial entender a diferenciação entre os provedores, tema sobre o qual há certa divergência. O MCI trata, no seu artigo 5º, incisos V e VII de dois tipos de provedores: os de acesso à internet e os que disponibilizam aplicações na rede. O provedor de conexão à internet (também chamado de provedor de acesso) é o que oferece aos usuários os meios para que se realize a conexão à Internet. Como somente fornece o meio, não é da sua função propiciar ferramentas para a geração de conteúdo. O provedor de aplicação, por sua vez, é “a pessoa que fornece um conjunto de funcionalidades que são acessadas por meio de um terminal conectado à internet” (SOUZA, Carlos Affonso; LEMOS, Ronaldo; BOTTINO, Celina (Coord.). *Marco Civil da Internet: jurisprudência comentada*. São Paulo: Revista dos Tribunais, 2017, p. 96). Ainda segundo os autores, podem englobar os provedores de conteúdo (como redes sociais, aplicativos de mensagens e plataformas de vídeos) e de hospedagem (que são os responsáveis por armazenar dados de terceiros). Atualmente, encontra-se em discussão a constitucionalidade do artigo 19 do MCI. Há repercussão geral no tema 987, cujo *leading case* é o RE 1.037.396/SP, relatado pelo Ministro Dias Toffoli (BRASIL. *Supremo Tribunal Federal (STF)*. Tema 987. [s.d.]. Disponível em: <http://stf.jus.br/portal/jurisprudenciaRepercussao/verAndamentoProcesso.asp?incidente=5160549&numeroProcesso=1037396&classeProcesso=RE&numeroTema=987>. Acesso em: 09 jan. 2022).

¹⁴ Disponível em: <https://pt-br.facebook.com/help/instagram/581066165581870>. Acesso em: 15 dez. 2022.

Nesse item também são trazidas informações sobre o encerramento da conta do usuário e da exclusão dos conteúdos por ele publicados.

Ao tratar da remoção de conteúdo, há direcionamento para as Diretrizes da Comunidade¹⁵. Inicialmente há ênfase para a atuação da rede na temática da COVID-19, afirmando que há o objetivo de evitar desinformações sobre a doença e as vacinas. Nesse caso, há uma política exclusiva da plataforma sobre o tema¹⁶.

De modo geral, as Diretrizes da Comunidade afirmam que o *Instagram* quer ser um local seguro para a expressão, aberto e seguro para todos. Segundo o texto, os usuários podem auxiliar nesse sentido ao publicar fotos e vídeos que estejam de acordo com a lei (proibição de exaltação ao terrorismo, crime organizado e grupos organizados de propagação de ódio, por exemplo), que respeitem a todos, que não sejam *spam* e que não contenham nudez (excetuando casos de situações relacionadas à saúde, atos de protesto e nudez em imagens e esculturas). Há também aviso específico proibindo conteúdos que violem os direitos de propriedade intelectual, que estimulem a automutilação e que expressem violência explícita.

No que se refere ao *Facebook*, a rede apresenta nos seus Padrões da Comunidade¹⁷ algumas questões sobre conteúdos proibidos e permitidos. Na introdução, há a informação de que a Meta considera importante o papel de eliminar de seus serviços qualquer tipo de abuso. Para isso, os Padrões são elaborados com base no feedback das pessoas e orientações de especialistas de área como tecnologia, segurança pública e direitos humanos.

Além disso, há um “compromisso com a voz” da rede, que diz possuir como prioridade a expressão. No entanto, tal direito poderia ser limitado quando há conflitos com questões de autenticidade (falsificação de identidade, por exemplo), segurança (risco de danos à segurança física das pessoas), privacidade (proteção das informações pessoais) e dignidade (proteção contra desrespeito, assédio e difamações).

Na primeira seção sobre “violência e comportamento criminoso” coíbem-se atividades de violência, incitação e organizações e indivíduos perigosos, por exemplo. A segunda seção trata sobre “segurança” e traz questões sobre suicídio, exploração sexual, *bullying*, assédio e privacidade. Na terceira, sobre “conteúdo questionável”, é esclarecida a proibição a discurso de ódio, violência e nudez, dentre outras. Na quarta parte, o aplicativo esclarece sobre “integridade e autenticidade”, explicando sobre integridade da conta, spam, desinformação e transformação da

¹⁵ Disponível em: https://help.instagram.com/477434105621119?ref=igtos&helpref=faq_content. Acesso em: 15 dez. 2022.

¹⁶ Disponível em: https://help.instagram.com/697825587576762/?helpref=uf_share. Acesso em: 15 dez. 2022.

¹⁷ Disponível em: https://pt-br.facebook.com/communitystandards/content_related_requests. Acesso em: 16 dez. 2022.

página em memorial. A quinta seção traz sobre o “respeito à propriedade intelectual”. Por fim, a seção sobre “solicitação e decisões relativas a conteúdo” traz duas políticas: sobre solicitações de usuários (remoção de conta própria, remoção de conta de familiar falecido ou remoção de conta de usuário incapacitado) e proteção adicional de menores de idade.

Após a análise desses termos e políticas, percebe-se que são estabelecidos critérios, ainda que não exaustivos, sobre o que é permitido e o que é proibido em cada plataforma. Embora exista certo vínculo com as definições jurídicas, essas determinações são eminentemente privadas, até mesmo porque essas plataformas costumam atuar em diversos países, com sistemas jurídicos distintos.

Retomando as fases da atividade de moderação de conteúdo pelas plataformas, em um segundo momento os moderadores escolhidos pela plataforma tomam as decisões sobre o conteúdo. Essa atividade pode ocorrer automaticamente (por algoritmos, por exemplo) ou por atividade de avaliação humana. Nesse caso normalmente há a denúncia e decide-se sobre a violação ou não das regras da plataforma. É comum que os usuários afetados sejam notificados.

Por fim, um terceiro momento possível na moderação de conteúdo é justamente aquele no qual se insere a atuação do *Oversight Board*, ou seja, a possibilidade de supervisionar e rever as decisões da plataforma. A ideia é trazer maior escrutínio e transparência para o processo de tomada de decisões. Percebe-se que, em alguns casos, é complexo analisar como os critérios postos são interpretados e aplicados, por serem gerais ou abertos.

Não se deve deixar de observar, no entanto, a crítica sobre a tecnorregulação, que se sobrepõe ao Estado Democrático de Direito, segundo Eduardo Magrani. O autor salienta que a filtragem e retirada de conteúdo ocorrem de maneira automatizadas e invisíveis, sem haver a responsabilização perante o usuário quando essa retirada for ilegal ou desmotivada. O perigo, nesse caso, é que fiquemos sujeitos às regras da tecnologia e não ao Estado de Direito (MAGRANI, 2019, p. 253-254).

Nesse sentido, mesmo representantes do *Facebook* concordam. Nick Clegg já se manifestou (CLEGG, 2021) afirmando que as pessoas ficariam mais confortáveis com a maior transparência das plataformas. Sendo assim, menciona que a internet precisa de regras estipuladas por instituições eleitas de maneira democrática. Para Clegg, o *Facebook* possui *standards* sobre os conteúdos aceitáveis, mas a definição da linha do aceitável é questão de debate público, o qual deve ser legitimado. Assim, as empresas privadas não deveriam, sozinhas, tomar decisões sobre o que é ou não aceitável. Porém, na ausência de leis, as

decisões devem ser tomadas. Nesse sentido, menciona como decisão positiva a criação do *Oversight Board*.

4 O OVERSIGHT BOARD

Sobre o *Oversight Board*, algumas questões essenciais foram analisadas neste artigo. Primeiramente, procurou-se compreender se há fontes seguras para analisar a sua atuação na moderação de conteúdo da Meta. Após, buscou-se visualizar como ele é composto e como é seu funcionamento. O lapso temporal determinado foram os anos de 2021 e 2022, com análise das decisões e pareceres publicados até dezembro de 2022. A análise busca responder questões sobre a objetividade de sua atuação, sua transparência, a existência de eventual auditoria externa, a visibilidade de seus critérios de decisão e, de maneira mais geral, se sua atuação é apta para melhorar a transparência, a confiabilidade e assim abrir a caixa-preta da Meta, no que se refere a moderação de conteúdo.

O primeiro ponto que deve ser mencionado é a existência de *site*¹⁸, público e facilmente acessível, que busca disponibilizar conteúdos e decisões sobre o *Oversight Board*, o Comitê de Supervisão. No *site* é possível escolher o idioma de visualização dos conteúdos disponibilizados, havendo a versão em português. Logo na página inicial, o Comitê esclarece seu principal objetivo, que seria “garantir o respeito à liberdade de expressão por meio do julgamento independente”. Percebe-se, dessa forma, a essencialidade do direito à liberdade de expressão, inclusive na internet.

O *site* disponibiliza informações sobre o seu processo de formação, incluindo as consultas variadas em diversos países. Há, ainda, linha do tempo demonstrando o progresso na sua formação. A primeira etapa foi em novembro de 2018, quando foi estabelecido o conceito do Comitê, e o *Facebook* compartilhou a ideia desse órgão, com julgamento independente sobre as tomadas de decisões mais complexas sobre conteúdo. O Estatuto¹⁹ foi publicado, em sua versão final, em setembro de 2019. E o Comitê iniciou suas operações em 2020.

Ainda na página inicial, o Comitê menciona sobre o motivo do seu desenvolvimento, esclarecendo que a comunidade do *Facebook* possui mais de dois bilhões de pessoas, de modo

¹⁸ Disponível em: <https://oversightboard.com/>. Acesso em: 18 dez. 2022.

¹⁹ O Estatuto do Comitê de Supervisão está disponível em: <https://oversightboard.com/governance/>. Acesso em: 18 dez. 2022.

que a empresa não deve ser a única responsável por tomar decisões, até mesmo porque elas afetam a liberdade de expressão e a segurança *online* dos indivíduos. Assim, sua criação ocorreu para auxiliar o *Facebook* a definir o que deve ser removido, o que deve ser permitido e o porquê. Percebe-se assim, de certa forma, a tentativa de melhorar o maior escrutínio público sobre as decisões da plataforma, percepção de integrantes do *Facebook* como Nick Clegg, como visto.

Assim, as deliberações do *Oversight Board* são vinculantes e irão manter ou reverter as decisões sobre conteúdo. Desse modo, o *Facebook* se compromete a implementá-las, excetuados os casos em que essa aceitação se configure como violação da lei. Não é esclarecido nesse ponto qual lei será analisada, já que o *Facebook* é empresa com atuação global, sendo aplicada em vários países, com legislações distintas. Assim, não se sabe se a lei considerada será a dos Estados Unidos, país onde o *Facebook* tem sede, ou o país no qual o conteúdo foi publicado.

Ainda na página inicial, ao tratar da finalidade do Comitê, é novamente esclarecido que seu objetivo é promover a liberdade de expressão. Nesse ponto, é mencionado que também serão realizadas análises de conteúdo no *Instagram*.

No que se refere ao Comitê, há informação na página inicial de que, quando completo, ele será composto por 40 membros de todo o mundo, com diversas formações acadêmicas e origens. Cada um dos membros terá mandato de três anos, pelo máximo de três mandatos (artigo 1.3 do Estatuto). Atualmente, o Comitê de Supervisão é composto por 23 membros²⁰. Dentre eles, são 12 homens e 11 mulheres, demonstrando a paridade de gênero.

Há seis representantes dos Estados Unidos da América, e um representante dos seguintes países: Iêmen, Reino Unido, Colômbia, Paquistão, Senegal e Gana, Dinamarca, Índia, Taiwan, Egito, Camarões e França, Austrália, Quênia, México, Hungria, Israel, Indonésia e Brasil. Assim totalizam-se os 23 membros, já que duas integrantes representam dois países (Afia Asantewaa Asare-Kyei representa Senegal e Gana e Julie Owono Camarões e França). São membros com experiências variadas, mas principalmente na área de direito, proteção de direitos humanos, liberdade de expressão e tecnologia.

Também na página inicial são mencionados os princípios do *Oversight Board*, que são: independência, autoridade, acessibilidade e transparência. A independência aparece como valor central do ente, já que essencial para que possa exercer sua função. Para isso, uma das previsões é

²⁰ Sobre o Comitê e sua formação, é disponibilizado uma página para conhece-lo, disponível em: <https://oversightboard.com/meet-the-board/>. Acesso em: 18 dez. 2022.

a ausência de conflito de interesses entre os membros e o seu processo de julgamento, além de experiência em deliberação e conhecimento sobre os temas relevantes para as decisões, segundo o artigo 1.2 do Estatuto. Outra forma para garantir a independência é a questão financeira, sendo esclarecido que o financiamento do Comitê é realizado por um *trust*²¹ independente, com apoio de uma empresa também independente, sem vínculo com o *Facebook*.

A autoridade ocorre porque, como visto, as decisões do Comitê são vinculantes, salvo violação de lei. A mesma noção também está expressa no artigo 4 do Estatuto. Também é esclarecido que o comitê pode emitir recomendações sobre as políticas para a empresa Meta.

A transparência é efetivada com a divulgação pública das declarações sobre as decisões e justificativas, que são publicadas no mesmo *site*. Além disso, são previstos relatórios anuais sobre o seu trabalho.

Sobre a acessibilidade, algumas considerações são relevantes. Ela é garantida através da possibilidade de apelação²², para o Comitê, das decisões inicialmente realizadas pelo *Facebook* e *Instagram*. Haverá análise sobre a manutenção ou a reversão dessas decisões de conteúdo, visando determinar se elas estão de acordo com as políticas e valores declarados pela plataforma.

Após a apelação haverá a seleção dos casos que serão efetivamente investigados, já que nem todos terão essa oportunidade de revisão. Há disponibilização dos requisitos sobre o envio da apelação. Na parte do site sobre o processo de apelação, é informado que o Comitê recebe milhares de apelações por semana, selecionando apenas algumas por mês. É esclarecido ainda que são selecionadas as decisões de conteúdo mais difíceis, com relevância global e que são relevantes para a elaboração de políticas futuras.

O artigo 2 do Estatuto do Comitê esclarece que, na seleção dos casos analisados, “o comitê buscará considerar os casos com maior potencial para orientar futuras decisões e políticas”²³. Assim, os membros do Comitê possuem poder de escolha e decisão de quais solicitações serão analisadas.

Considera-se que as informações sobre essas escolhas são abstratas, ao considerar que os casos qualificados serão “difíceis”, com “relevância global” e capazes de auxiliar na elaboração de futuras políticas das plataformas. Como se percebe, tratam-se de expressões genéricas e subjetivas, sem parâmetros concretos sobre a definição realizada pelos membros do Comitê.

²¹ Sobre os *trustees*, o Comitê disponibiliza página específica sobre quem são e sua atuação. Disponível em: <https://oversightboard.com/governance/>. Acesso em: 18 dez. 2022.

²² O processo de apelações é explicado no seguinte *link*: <https://oversightboard.com/appeals-process/>. Acesso em: 18 dez. 2022.

²³ Disponível em: <https://oversightboard.com/governance/>. Acesso em: 19 dez. 2022.

4 BREVE INVENTÁRIO DAS DECISÕES DO *OVERSIGHT BOARD*

As decisões do Comitê são disponibilizadas publicamente²⁴, como já mencionado. Apesar das decisões serem públicas, há o cuidado de proteger a privacidade dos indivíduos²⁵ que publicaram ou que sejam afetados pelo conteúdo disponibilizado nas redes sociais.

Sobre as decisões do Comitê, em análise realizada em dezembro de 2022 haviam 33 decisões disponibilizadas. Há, na página, um quadro com informações iniciais sobre a manutenção ou reversão da decisão inicial da plataforma, informações sobre seu número, um título e resumo, qual rede social houve publicação do conteúdo, a matéria, o país e a data de publicação da decisão do Comitê. Além disso, há uma página específica para cada decisão do *Oversight Board*, com resumo do caso, explicação sobre ele, as principais conclusões e a decisão do Comitê. Após, apresenta-se a decisão completa sobre o caso. Os casos com a sigla “FB” no nome se referem a decisões de conteúdo do *Facebook*. A sigla “IG”, por sua vez, sinaliza que a plataforma cuja decisão é analisada é o *Instagram*.

Assim, menciona-se inicialmente o panorama sobre as 33 decisões encontradas. O lapso temporal analisado foi de dois anos, desde a primeira decisão do comitê, em janeiro de 2021, até a trigésima terceira decisão em 14 de dezembro do ano subsequente. Uma delas (2020-001-FB-UA²⁶) não foi analisada pelo Comitê. A indisponibilidade ocorreu porque a publicação foi excluída pelo próprio usuário que a divulgou, na qual estava o comentário objeto da discussão.

Foram vinte e três as decisões do Comitê no sentido de reverter as decisões iniciais das plataformas, seja o *Facebook* ou *Instagram*. O caso 2020-002-FB-UA²⁷ tratava de publicação com duas fotografias, de um bebê sírio de etnia curda. O texto acompanhado dizia que “muçulmanos devem ter algum problema psicológico ou mental”. A exclusão se baseou na regra de Discurso de Ódio. No entanto, o Comitê revogou a decisão do Facebook, por considerar que o conteúdo, embora ofensivo, não devia ser enquadrado como discurso de ódio.

²⁴ Disponível em: <https://oversightboard.com/decision/>. Acesso em: 19 dez. 2022.

²⁵ No item “Como protegemos a sua privacidade se o Comitê analisar o seu caso” é esclarecido que a decisão final é disponível publicamente no site. No entanto, deverá haver uma permissão do titular de dados para o compartilhamento de detalhes que facilitem a sua identidade. Disponível em: <https://oversightboard.com/appeals-process/>. Acesso em: 19 dez. 2022.

²⁶ Disponível em: <https://oversightboard.com/decision/FB-KBHZS8BL/>. Acesso em: 19 dez. 2022.

²⁷ Disponível em: <https://oversightboard.com/decision/FB-I2T6526K/>. Acesso em: 19 dez. 2022.

O caso 2020-004-IG-UA²⁸ foi o primeiro analisado sobre conteúdo postado no Brasil. Trata-se de caso no qual o conteúdo foi restaurado pelo *Facebook*, após a decisão do Comitê, por entender-se que a publicação não violava as regras sobre nudez e atividade sexual de adultos, mas sim representava campanha sobre o *Outubro Rosa* e o combate ao câncer de mama. O caso traz um alerta importante sobre a moderação automatizada, que pode violar direitos humanos. Por isso são importantes previsões como a do *Instagram*, já mencionada, que proíbe a divulgação de imagens de nudez, mas excetua os casos de situações relacionadas à saúde.

No caso 2020-005-FB-UA²⁹ houve a postagem de um usuário sobre uma citação atribuída a Joseph Goebbels, Ministro da Propaganda no governo nazista. A postagem foi enquadrada na regra de padrões sobre Organizações e Indivíduos Perigosos. Não havia fotos de Goebbels ou símbolos nazistas na publicação, e o usuário argumentou que tentava comparar Donald Trump ao regime nazista. O Comitê considerou que não houve apoio ao nazismo na publicação e que as regras não eram suficientemente claras aos usuários, o que viola os direitos humanos e a liberdade de expressão. Assim, revogou a decisão inicial.

A publicação de conteúdo sobre remédios como hidroxicloroquina e azitromicina, vinculando-os ao tratamento da COVID-19 foi analisada no caso 2020-006-FB-FBR³⁰. Inicialmente houve a exclusão pela plataforma, por contribuir para o risco de danos físicos eminentes, regra que integra o Padrão da Comunidade sobre Violência e Incitação. O Comitê considerou que as regras de desinformação e danos eminentes são vagas. Além disso, determinou que a empresa crie um Padrão da Comunidade claro sobre questões de desinformação sobre saúde.

A publicação de meme em um fórum para muçulmanos indianos, incluindo críticas ao presidente francês Emmanuel Macron, foi objeto de averiguação no caso 2020-007-FB-FBR³¹. O conteúdo foi inicialmente excluído por ameaça de violência, o que pode violar os Padrões sobre Violência e Incitação. A maioria do Comitê considerou que as incitações não são necessariamente violentas, e interpretou-as como críticas, decidindo pela restauração da postagem.

²⁸ Disponível em: <https://oversightboard.com/decision/IG-7THR3SI1/>. Acesso em: 19 dez. 2022.

²⁹ Disponível em: <https://oversightboard.com/decision/FB-2RDRCVQ/>. Acesso em: 19 dez. 2022.

³⁰ Disponível em: <https://oversightboard.com/decision/FB-XWJQBU9A/>. Acesso em: 19 dez. 2022.

³¹ Disponível em: <https://oversightboard.com/decision/FB-R9K87402/>. Acesso em: 19 dez. 2022.

No caso 2021-003-FB-UA³² anulou-se a remoção do *Facebook* sobre vídeo, na Índia, sobre questões religiosas, e o conteúdo foi restaurado. Nesse caso, houve denúncia de publicação que violaria o Padrão da Comunidade sobre Organizações e Indivíduos Perigosos. O *Facebook* não realizou análise humana, sob o argumento de redução temporária de pessoal pela pandemia de COVID-19. O Comitê criticou a ausência da análise e, além disso, incentivou que a empresa tome medidas para evitar o silenciamento de vozes de minorias religiosas.

Na decisão 2021-005-FB-UA³³ revogou-se a decisão de excluir conteúdo com base no Padrão de Discurso de Ódio. No caso, se tratava de um meme para conscientizar e condenar as falas do governo turco de negação do genocídio armênio, ao mesmo tempo em que as justifica. Para a maioria do Comitê, o caso se enquadra como sátira e, portanto, como exceção. A intenção seria justamente condenar ou aumentar a conscientização sobre o discurso de ódio, de modo que o conteúdo deveria permanecer na plataforma.

O caso 2021-004-FB-UA³⁴ trouxe a revogação da decisão do *Facebook* de remover o comentário de apoiador de Alexei Navalny, opositor russo, que chamou outro usuário de “robô covarde”. Entendeu o Comitê que, nesse caso, não se aplicava o Padrão da Comunidade sobre *Bullying* e Assédio, e que a manifestação estava de acordo com a liberdade de expressão.

No caso 2021-006-IG-UA³⁵ o conteúdo foi removido e depois restaurado pelo próprio *Facebook*. Se tratava de discussão das condições de confinamento de Abdullah Öcalan, membro fundador do PKK, a prisão na Ilha Imrali, na Turquia. Nesse caso, em razão da preocupação do Comitê, o *Facebook* argumentou que está trabalhando na atualização de políticas para permitir a discussão de direitos humanos de indivíduos classificados como perigosos. O conteúdo havia sido inicialmente removido pela regra do *Facebook* sobre organizações e indivíduos perigosos.

Também houve discussão sobre discurso de ódio no caso 2021-007-FB-UA. Porém, o Comitê entendeu que a publicação foi direcionada para o estado chinês (especificamente a política governamental chinesa em Hong Kong e em Mianmar), não as pessoas chinesas, e mandou restaurar a publicação.

No caso 2021-009-FB-UA³⁶ o *Facebook* inicialmente removeu o conteúdo, mas depois o restaurou, e o Comitê considerou que a restauração era a medida acertada. No caso, o conteúdo mencionava notícia sobre uma ameaça de violência das Brigadas Izz al-Din e al-

³² Disponível em: <https://oversightboard.com/decision/FB-H6OZKDS3/>. Acesso em: 19 dez. 2022.

³³ Disponível em: <https://oversightboard.com/decision/FB-RZL57QHJ/>. Acesso em: 19 dez. 2022.

³⁴ Disponível em: <https://oversightboard.com/decision/FB-6YHRXHZR/>. Acesso em: 19 dez. 2022.

³⁵ Disponível em: <https://oversightboard.com/decision/IG-I9DP23IB/>. Acesso em: 19 dez. 2022.

³⁶ Disponível em: <https://oversightboard.com/decision/FB-P93JPX02/>. Acesso em: 19 dez. 2022.

Qassam, ala militar do grupo palestino Hamas. O Comitê considerou que o tema é de interesse público e, como tal, não se justificaria a limitação da liberdade de expressão. Salienta-se que tal caso está incluído como “revertido” mas, no entanto, trata-se de concordância do Comitê com a segunda e final decisão do *Facebook*, no sentido de manter o conteúdo.

A revogação da decisão de remover vídeo de manifestantes na Colômbia que criticavam o presidente, Ivan Duque, foi objeto de análise no caso 2021-010-FB-UA³⁷. Uma das palavras mencionadas foi entendida como calúnia e, portanto, classificada como Discurso de Ódio. No entanto, o Comitê entendeu que há interesse público e permissão de conteúdo pelo valor jornalístico no caso.

O caso 2021-012-FB-UA³⁸ trata de mudança da exclusão inicial de conteúdo de artista indígena norte-americano, com base no Discurso de Ódio. Entendeu o Comitê que se trata de exceção, que objetiva aumentar a divulgação de crimes históricos contra indígenas na América do Norte.

Outro exemplo de publicação brasileira foi o caso 2021-013-IG-UA³⁹. Revogou-se a decisão da Meta de remover publicação sobre a bebida à base de plantas ayahuasca. Entendeu o Comitê que não há violação às regras do *Instagram*. Ademais, recomendou que as regras permitam as discussões sobre usos tradicionais ou religiosos de drogas não medicinais.

O caso 2021-015-FB-UA⁴⁰ menciona a decisão original da Meta de remover publicação do *Facebook* que pedia conselhos sobre consultas médicas, o diagnóstico e tratamento do transtorno do déficit de atenção com hiperatividade (TDAH) e o medicamento controlado *Adderall*. O Comitê considerou que era caso de aconselhamento, e não havia conexão entre o conteúdo e possibilidade de dano aos usuários.

Na decisão 2021-016-FB-FBR⁴¹ o Meta removeu publicação que continha situações de violência sexual contra menores de idade em denúncia de jornalista Sueca. O conteúdo foi restaurado pois não se tratava de caso de exploração sexual, mas sim da condenação de tal conduta e informação sobre tema de interesse público e social. Na descrição realizada não há identificação das vítimas ou dos criminosos.

O caso 2022-003-IG-UA⁴² traz a anulação da decisão da Meta de remover publicação do *Instagram* que continha fotos de palavras em árabe que poderiam ser consideradas

³⁷ Disponível em: <https://oversightboard.com/decision/FB-E5M6QZGA/>. Acesso em: 19 dez. 2022.

³⁸ Disponível em: <https://oversightboard.com/decision/FB-L1LANIA7/>. Acesso em: 19 dez. 2022.

³⁹ Disponível em: <https://oversightboard.com/decision/IG-0U6FLA5B/>. Acesso em: 19 dez. 2022.

⁴⁰ Disponível em: <https://oversightboard.com/decision/FB-Q72FD6YL/>. Acesso em: 19 dez. 2022.

⁴¹ Disponível em <https://oversightboard.com/decision/FB-P9PR9RSA/>. Acesso em: 19 dez. 2022.

⁴² Disponível em: <https://oversightboard.com/decision/IG-2PJ00L4T/>. Acesso em: 19 dez. 2022.

depreciativas para homens com “trejeitos afeminados”. A publicação não poderia ser considerada discurso de ódio porque tratava-se de tentativa de ressignificar os termos ofensivos. Logo, o *Board* decidiu que, nesse caso, o discurso de ódio é somente citado para condenar e conscientizar sobre o assunto, discurso que não deve ser proibido.

O caso 2022-001-FB-UA⁴³ menciona a manutenção do *Facebook* de publicação com a versão editada de uma animação que comparava sérvios com ratos. Mesmo com as denúncias e revisão humana adicional, o conteúdo foi mantido. Porém, tal decisão foi anulada pelo Comitê, que considerou que o conteúdo continha ataque contra pessoas por sua etnia, comparações degradantes e intolerantes e celebração de discriminação. Portanto, trata-se de violação aos Padrões de Comunidade sobre discurso de ódio e violência e incitação na plataforma.

A ilustração que retratava violência policial na Colômbia, no caso 2022-004-FB-UA⁴⁴, foi originalmente excluída pela Meta. O motivo foi a correspondência da publicação com uma das imagens do banco de dados do Serviço de Correspondência de Mídias. Nesse caso, o banco identifica e remove de maneira automática imagens que sejam identificadas como violadoras das regras da empresa. A decisão foi revogada pelo Comitê, que considerou que não havia violação das regras, restaurando a publicação. O Comitê recomendou ainda que a Meta melhore os seus procedimentos de remoção e os conteúdos desses bancos de dados.

No caso 2022-005-FB-UA⁴⁵ a Meta inicialmente removeu publicação do *Facebook* de veículo de notícias que relatava anúncio positivo do regime Talibã no Afeganistão, relacionado com a educação de mulheres e meninas no país. O Comitê considerou que não se tratava de violação aos Padrões no que se refere a organizações e indivíduos perigosos, mas sim um caso de liberdade de expressão dos usuários sobre notícias referentes a regimes terroristas.

A decisão de remover publicação no *Facebook* que comparava o exército russo na Ucrânia com os nazistas e que citava poema que pedia o extermínio de fascistas foi anulada pelo Comitê no caso 2022-008-FB-UA⁴⁶. Na publicação também havia, aparentemente, imagem de um cadáver, removido por violar a política de conteúdo violento e explícito. O Comitê considerou a importância do contexto e que, nesse caso, havia descrição de momentos históricos de violência, sem encorajar violência. Foi considerado também o caso específico, de reconhecimento internacional da ilegalidade da invasão russa na Ucrânia.

⁴³ Disponível em: <https://oversightboard.com/decision/FB-JRQ1XP2M/>. Acesso em: 19 dez. 2022.

⁴⁴ Disponível em: <https://oversightboard.com/decision/FB-I964KKM6/>. Acesso em: 19 dez. 2022.

⁴⁵ Disponível em: <https://oversightboard.com/decision/FB-U2HHA647/>. Acesso em: 21 dez. 2022.

⁴⁶ Disponível em: <https://oversightboard.com/decision/FB-MBGOTVN8/>. Acesso em: 21 dez. 2022.

O *Instagram* decidiu remover videoclipe do gênero musical *drill*, e o Comitê decidiu restaurá-lo no caso 2022-007-IG-UA⁴⁷. A remoção ocorreu após solicitação da Polícia Metropolitana, no Reino Unido, que se preocupou que a música causasse mais violência ao fazer referência a um tiroteio de 2017. O Comitê considerou que é necessário melhorar o devido processo e a transparência nas relações da rede social com a polícia, e que não havia violação à Política de violência e incitação.

O último caso de reversão da decisão é o 2022-011-IG-UA⁴⁸, no qual a Meta removeu vídeo do *Instagram* que mostrava ataque terrorista em Igreja na Nigéria. Nesse caso também houve correspondência com bancos de dados do Serviço de Correspondência de Mídias, que removeu o conteúdo. Após análise do Comitê, a publicação foi restaurada, com aviso que protege a privacidade das vítimas. Porém, considerou-se que era importante permitir a discussão e proteger a liberdade de expressão, especialmente em contexto social de ataques terroristas, como ocorre na Nigéria na atualidade.

Passa-se agora ao exame das decisões do Comitê pela manutenção das escolhas inicialmente realizadas pelas plataformas, que totalizam nove. No caso 2020-003-FB-UA⁴⁹ decidiu-se pela manutenção da remoção do conteúdo, por violar as regras da comunidade sobre Discurso de Ódio. Tratava-se de publicação que insultava e depreciava um grupo de pessoas, no caso, os azerbaijanos. Por sua vez, o caso 2021-002-FB-UA⁵⁰ trouxe a decisão pela manutenção de exclusão de caricaturas de pessoas negras na forma de *blackface*, também por violar o Padrão de Comunidade sobre Discurso de Ódio.

O caso 2021-001-FB-FBR⁵¹ traz decisão bastante conhecida e discutida das redes sociais e do próprio Comitê. Decidiu-se pela continuidade da restrição do ex-presidente dos Estados Unidos, Donald Trump, às suas contas no *Facebook* e *Instagram*. Porém, o Comitê entendeu que não é apropriada a imposição de penalidades com prazo indeterminado e fora dos Padrões. Além disso, recomendou que o *Facebook* desenvolva políticas claras sobre a segurança pública e o respeito à liberdade de expressão.

Outro exemplo brasileiro é a decisão 2021-008-FB-FBR⁵², que menciona publicação de um conselho médico estadual afirmando que o confinamento é ineficaz como medida preventiva da COVID-19, e que foi condenado pela Organização Mundial da Saúde (OMS). A

⁴⁷ Disponível em: <https://oversightboard.com/decision/IG-PT5WRTLW/>. Acesso em: 21 dez. 2022.

⁴⁸ Disponível em: <https://oversightboard.com/decision/IG-OZNR5J1Z/>. Acesso em: 21 dez. 2022.

⁴⁹ Disponível em: <https://oversightboard.com/decision/FB-QBJDASCV/>. Acesso em: 19 dez. 2022.

⁵⁰ Disponível em: <https://oversightboard.com/decision/FB-S6NRTDAJ/>. Acesso em: 19 dez. 2022.

⁵¹ Disponível em: <https://oversightboard.com/decision/FB-691QAMHJ/>. Acesso em: 19 dez. 2022.

⁵² Disponível em: <https://oversightboard.com/decision/FB-B6NGYREK/>. Acesso em: 19 dez. 2022.

decisão do *Facebook* de manter a publicação foi corroborada pelo Comitê. Considerou-se que o conteúdo não criava riscos de danos iminentes, podendo permanecer na plataforma. O *Oversight Board*, no entanto, enfatizou a importância de outras medidas para combater a disseminação de desinformação sobre a COVID-19, além da remoção.

No caso 2021-011-FB-UA⁵³ manteve-se a remoção, com base no Discurso de Ódio, de publicação que discutia a sociedade sul-africana. Entendeu o ente em sua decisão que a publicação continha calúnias degradantes, excludentes e prejudiciais, já que se direcionavam para as pessoas em razão de sua raça, etnia e/ou nacionalidade.

O caso 2021-014-FB-UA trouxe a decisão do Comitê de manter decisão da Meta de remoção de publicação. Alega-se, no caso, que a Frente de Libertação do Povo Tigré (TPLF), tinha cometido crimes na região de Amara, Etiópia. Salienta-se que, nesse caso, houve desencontro de decisões, já que a Meta restaurou o conteúdo após a apelação do usuário. O Comitê decidiu que o conteúdo deve ser retirado novamente, por violar os Padrões sobre Violência e Incitação.

Um vídeo explícito de violência contra civil no Sudão foi restaurado pela Meta, e a decisão foi mantida pelo Comitê no caso 2022-002-FB-MR⁵⁴. A publicação, inicialmente removida por conter conteúdo explícito e violento, sofreu apelação pelo usuário e a Meta considerou que o conteúdo era de valor jornalístico, restaurando o conteúdo com uma tela de aviso no vídeo, impedindo o acesso de menores. O Comitê considerou que o material aumentava a conscientização sobre questões de direitos humanos, tendo relevante interesse público. Fez também recomendação para a Meta de autorização de vídeos com cadáveres ou pessoas feridas, quando forem compartilhados para conscientizar ou retratar abusos.

No caso 2022-006-FB-MR⁵⁵ a Meta removeu publicação com ameaça de violência no conflito na Etiópia, e o Comitê de Supervisão manteve a remoção. Considerou-se que havia violação do Padrão da Comunidade da Meta sobre violência e incitação. Segundo o Comitê, a exclusão estaria de acordo com a responsabilidade empresarial de proteção dos direitos humanos. A manutenção da publicação tinha risco de provocar mais violência no país, em razão do contexto de conflitos na Etiópia.

Por fim, o último caso publicado em 2022 é de manutenção da decisão da Meta de restaurar vídeo publicado no *Instagram* que mostrava assédio sexual de um grupo de homens

⁵³ Disponível em: <https://oversightboard.com/decision/FB-TYE2766G/>. Acesso em: 19 dez. 2022.

⁵⁴ Disponível em: <https://oversightboard.com/decision/FB-AP0NSBVC/>. Acesso em: 21 dez. 2022.

⁵⁵ Disponível em: <https://oversightboard.com/decision/FB-E1154YLY/>. Acesso em: 21 dez. 2022.

na Índia. No caso 2022-012-IG-MR⁵⁶ considerou-se que deve haver tolerância da empresa com conteúdo que possua interesse jornalístico. No vídeo a vítima não é identificada e não há nudez, mas houve inserção de tela de aviso que impede o acesso a menores de 18 anos. O Comitê considerou que a publicação serviu para conscientização e possui utilidade pública.

5 AVALIAÇÃO PRELIMINAR

Das 33 (trinta e três) decisões ocorridas até o final de 2022, 1 (uma) não pode ser analisada, pela exclusão do conteúdo pelo próprio usuário. Das demais, 23 (vinte e três) foram pela revogação da decisão inicial (do *Facebook* ou *Instagram*). As 9 (nove) demais foram pela manutenção da decisão originária das plataformas. Apurou-se também que três decisões (2020-004-IG-UA; 2021-008-FB-FBR e 2021-013-IG-UA) são de conteúdos originariamente publicados no Brasil.

Os diagnósticos aqui realizados são iniciais, e, por isso, não há aprofundamento sobre os argumentos discutidos nas decisões do *Oversight Board*. Não se objetiva exaurir cada decisão, mas visualizar algumas das alegações principais sustentadas pelo Comitê e compreender os seus procedimentos e funcionamento.

Pela análise realizada pode-se visualizar a grande relevância do direito à liberdade de expressão pelo Comitê. De acordo com a análise realizada, dos 33 casos analisados, um tem a análise inviabilizada pela exclusão, um é da restrição a contas do ex-presidente Donald Trump e somente em 6 defendeu-se a exclusão do conteúdo postado, após as decisões da Meta e do Comitê.

Percebe-se também uma tentativa de compatibilizar a liberdade de expressão com outros direitos relevantes, a exemplo da privacidade das vítimas (casos 2021-016-FB-FBR e 2022-012-IG-MR) e proteção de menores com inserção de tela de aviso (casos 2022-002-FB-MR e 2022-012-IG-MR).

Além disso, há análise, por parte do Comitê, do contexto histórico e social do país, considerando se o assunto da publicação é relevante para discussão, especialmente para minorias, como nos exemplos dos casos 2021-005-FB-UA e 2022-003-IG-UA. Também há análise sobre o motivo que justifica a publicação do conteúdo violento, ou seja, se é uma

⁵⁶ Disponível em: <https://oversightboard.com/decision/IG-KFLY3526/>. Acesso em: 21 dez. 2022.

denúncia, esclarecimento ou sadismo (casos 2021-011-FB-UA, 2022-012-IG-MR, 2022-011-IG-UA e 2022-004-FB-UA, dentre outros).

Há um interesse na preservação do interesse jornalístico (casos 2021-010-FB-UA e 2022-005, FB-UA, por exemplo), da produção cultural (casos 2021-012-FB-UA e 2022-007-IG-MR) e da circulação de informações relevantes para o público, para conscientização (caso 2021-016-FB-FBR). Além disso, cita-se como relevante a preocupação do Comitê em não incitar violência em conflitos já estabelecidos (caso 2022-006-FB-MR), por considerar uma limitação aceita à liberdade de expressão.

Após tais análises, sustenta-se que o Comitê possui grande importância para trazer maior justiça e objetividade às decisões da Meta. Em sentido mais imediato, é possível haver a revisão humana, por pessoas capacitadas para tal função. Além disso, o Comitê é vantajoso para solucionar problemas futuros sobre moderação de conteúdo. Havendo indefinições ou regras muito abertas, por exemplo, o ente também age trazendo recomendações à Meta, para que sejam aperfeiçoadas as regras de utilização da plataforma já vigentes. Isso ocorreu, por exemplo, em casos sobre desinformação sobre a saúde (2020-006-BR-FBR e 2021-008-FB-FBR) e na proposta de novas medidas para evitar o selecionamento de minorias religiosas (2021-003-FB-UA).

O Comitê também atua na publicação de pareceres, após consulta da Meta, havendo dois disponíveis no site em dezembro de 2022. No primeiro parecer consultivo, PAO-2021-01⁵⁷, há recomendações do *Board* sobre o compartilhamento de informações residenciais privadas consideradas “disponíveis publicamente”. Considerou-se que tal conduta poderia colocar em risco a segurança e privacidade dos indivíduos.

A segunda opinião consultiva, PAO-2021-02⁵⁸, verifica o programa de verificação cruzada da Meta. Essa ocorreria a partir de uma lista de usuários para os quais as publicações inicialmente identificadas como violadoras passariam por uma segunda camada de análise, humana. O Comitê considerou que tal conduta traz questionamentos, a exemplo do tratamento desigual dos usuários, remoção atrasada de conteúdo violador e falta de transparência sobre o funcionamento da verificação cruzada.

Desse modo, a criação do *Oversight Board* parece cumprir a exigência de realização de um debate na esfera pública que envolva sociedade, Estado e empresas (MAGRANI, 2019, p. 259), sobre os conteúdos disponibilizados e as eventuais exclusões nas redes sociais.

⁵⁷ Disponível em: <https://oversightboard.com/decision/PAO-2021-01/>. Acesso em: 22 dez. 2022.

⁵⁸ Disponível em: <https://oversightboard.com/decision/PAO-NR730OFI/>. Acesso em: 22 dez. 2022.

Sustenta-se também que o ente cumpre com a preocupação da inteligibilidade de Frank Pasquale. As decisões do Comitê são disponíveis publicamente. No entanto, além da mera transparência, os motivos das decisões são explicados de maneira acessível, facilitando a sua compreensão mesmo para o público leigo.

6 CONSIDERAÇÕES FINAIS

O presente artigo discutiu a atuação do *Oversight Board*, da empresa Meta, no intuito de propiciar maior clareza e justiça às decisões de plataformas como *Facebook* e *Instagram* em decisões, por vezes complexas, sobre moderação de conteúdo. Inicialmente realizou-se análise sobre o desenvolvimento das redes sociais como forma de comunicação entre os indivíduos. Assim, aumentou-se também a publicação de conteúdos diversos nessas plataformas.

Como visto, o direito à liberdade de expressão ganha novos contornos na internet e nas redes sociais, com variados conteúdos sendo disponibilizados de maneira livre, a baixo custo, e com possibilidade de atingir número indeterminado de pessoas em pouco tempo. Desse modo, nem sempre as definições sobre quais conteúdos podem ser mantidos nas redes sociais são fáceis. Na ausência de legislações definitivas sobre o assunto, o que ocorre, em muitos casos, é a regulação pela própria plataforma, através de regras próprias sobre moderação de conteúdo incluídas em instrumentos como os Termos de Uso.

Foi analisado o papel do *Oversight Board* como instrumento de transparência digital da empresa Meta. Além disso, descreveu-se seu funcionamento e composição. Apesar de ter sido criado para auxiliar o *Facebook*, o Comitê se coloca como independente, esclarecendo que seu julgamento também é autônomo, baseando-se no apoio à liberdade de expressão das pessoas.

Algumas incertezas foram visualizadas na atuação do Comitê. Inicialmente mencionou-se a dificuldade de visualização sobre quais casos serão revistos por ele. Não há - ou, ao menos, não são disponibilizados de maneira pública para a sociedade - critérios objetivos que embasem a escolha dos casos que serão efetivamente revisados. Além disso, não foram informados os mecanismos de cumprimento forçado das suas decisões, que são obrigatórias, pela Meta. Também não são mencionadas informações sobre a possibilidade de revisões de suas decisões ou sobre a realização de auditorias externas. Considera-se que essas questões ainda necessitam de mais esclarecimentos. Por fim, sobre a sua composição, há atualmente seis membros dos Estados Unidos da América, dentre vinte e três membros totais. A

existência de 26% dos membros de um só país parece desprestigiar a diversidade, tão essencial ao Comitê.

No entanto, outras questões são extremamente positivas, como a presença de *site* público, em vários idiomas e de fácil acesso. Trata-se, ademais, de fonte segura para analisar sua composição e seu funcionamento. Os critérios de decisão são visíveis e apresentados de maneira objetiva. Sendo assim, considera-se que a sua atuação melhora a transparência e a confiabilidade da empresa Meta. Os mecanismos de funcionamento dos algoritmos de decisão continuam não disponíveis, mas a questões sobre moderação de conteúdo torna-se muito mais sujeita ao escrutínio público. Cumpre-se também com ideia de Pasquale da necessidade de transparência e inteligibilidade. As decisões são possíveis de compreender, mesmo para público leigo, já que as questões jurídicas não são centrais nos argumentos das decisões.

O *Oversight Board* possui um efeito mais direto e imediato, de tornar mais públicos, transparentes e acessíveis os mecanismos de moderação de conteúdo das redes sociais. Desse modo, abre-se, de certa maneira, a *Black Box* dessas redes. Em um segundo momento, mais a longo prazo, as recomendações feitas pelo Comitê também são extremamente úteis para propiciar normas de utilização das plataformas mais claras e objetivas aos usuários no futuro.

Ademais, à guisa de conclusão, o fomento à transparência e a devida publicidade das decisões e de seus fundamentos, a perspectiva de análise colegiada que atenua idiossincrasias e subjetividades censoras e a submissão ao debate público da motivação da moderação de conteúdo promovidas no modelo de *Oversight Board* são guias principiológicos importantes a serem observados em eventual regulamentação legislativa da matéria.

REFERÊNCIAS

ALBUQUERQUE, Karol. *Facebook* foi um dos protagonistas da invasão ao Capitólio, nos EUA. 22 out. 2021. *Olhar Digital*. Disponível em: <https://olhardigital.com.br/2021/10/22/internet-e-redes-sociais/facebook-foi-um-dos-protagonistas-da-invasao-ao-capitolio-nos-eua/>. Acesso em: 10 fev. 2022.

ANDRADE, André Gustavo Corrêa de. *Liberdade de expressão em tempos de cólera*. Rio de Janeiro: GZ, 2020.

APRESENTANDO Meta: uma empresa de tecnologia digital. 28 out. 2021. *Facebook*. Disponível em: <https://about.fb.com/br/news/2021/10/apresentando-meta-uma-empresa-de-tecnologia-social/>. Acesso em: 09 jan. 2022.

BINENBOJM, Gustavo. *Liberdade igual: o que é e por que importa*. Rio de Janeiro: História Real, 2020.

BRANCO, Sérgio. *Memória e esquecimento na internet*. Porto Alegre: Arquipélago Editorial, 2017.

BRASIL. *Supremo Tribunal Federal (STF)*. Tema 987. [s.d.]. Disponível em: <http://stf.jus.br/portal/jurisprudenciaRepercussao/verAndamentoProcesso.asp?incidente=5160549&numeroProcesso=1037396&classeProcesso=RE&numeroTema=987>. Acesso em: 09 jan. 2022.

CLEGG, Nick. *You and the Algorithm: It Takes Two to Tango*. 31 mar. 2021. Disponível em: <https://nickclegg.medium.com/you-and-the-algorithm-it-takes-two-to-tango-7722b19aa1c2>. Acesso em: 15 abr. 2021.

DIRETRIZES da Comunidade do *Instagram*. Meta, [s.l.], 2022. Disponível em: <https://help.instagram.com/477434105621119?ref=igtos>. Acesso em: 10 fev. 2022.

FACEBOOK (FBOK34), agora Meta, perde mais de US\$ 230 bi em valor de mercado em pior sessão de sua história. 03 fev. 2022. *InfoMoney*. Disponível em: <https://www.infomoney.com.br/mercados/acoes-do-facebook-fbok34-agora-meta-desabam-22-apos-balanco-com-numeros-abaixo-do-esperado/>. Acesso em: 09 fev. 2022.

GUSTIN, Miracy Barbosa de Sousa; DIAS, Maria Tereza Fonseca. *(Re)pensando a pesquisa jurídica: teoria e prática*. 4. ed. Belo Horizonte: Del Rey, 2013.

KELLER, Clara Iglesias. *Regulação nacional de serviços na internet: exceção, legitimidade e o papel do Estado*. Rio de Janeiro: Lumen Juris, 2019.

MAGRANI, Eduardo. *Entre dados e robôs: ética e privacidade na era da hiperconectividade*. 2. ed. Porto Alegre: Arquipélago Editorial, 2019.

MORAES, Maria Celina Bodin de; KONDER, Carlos Nelson. *Dilemas de Direito Civil Constitucional: casos e decisões sobre os novos desafios para a tutela da pessoa humana nas relações existenciais*. Rio de Janeiro: Renovar, 2012.

Oversight Board. Disponível em: <https://oversightboard.com/>. Acesso em: 07 jan. 2022.

PADRÕES da Comunidade do *Facebook*. Meta, [s.l.], 2022. Disponível em: https://pt-br.facebook.com/communitystandards/content_related_requests. Acesso em: 10 fev. 2022.

PASQUALE, Frank. *The Black Box Society: the secret algorithms that control money and information*. Cambridge: Harvard University Press, 2016.

PROTEÇÕES e Atualizações da Política sobre COVID-19 e Vacinas. Meta, [s.l.], 2022. Disponível em: https://help.instagram.com/697825587576762/?helpref=uf_share. Acesso em: 15 dez. 2022.

REGINATO, Andréa Depieri de A. Uma introdução à pesquisa documental. In: MACHADO, Máira Rocha (Org.). *Pesquisar empiricamente o direito*. São Paulo: Rede de Estudos Empíricos em Direito, 2017, p. 189-224.

ROBL FILHO, Ilton Norberto; MARRAFON, Marco Aurélio; PANSIERI, Flávio. Constitucionalismo como salvaguarda do Estado de Direito: crítica ao (ciber) populismo autoritário e a necessária reengenharia constitucional. *Revista do Instituto de Hermenêutica Jurídica: RIHJ.*, v. 18, 2020.

RODOTÀ, Stefano. *Il diritto di avere diritti*. Bari: Laterza & Figli, 2012.

SCHREIBER, Anderson. Direito e mídia. In: SCHREIBER, Anderson (Coord.). *Direito e Mídia*. São Paulo: Atlas, 2013.

SEISDEDOS, Iker. A informante que levou o *Facebook* à sua pior crise existencial. 10 out. 2021. *El País Brasil*. Disponível em: <https://brasil.elpais.com/tecnologia/2021-10-10/a-informante-que-levou-o-facebook-a-sua-pior-crise-existencial.html>. Acesso em: 10 fev. 2022.

SILVA, Gabriel Souza de. *O caso Cambridge Analytica*: distorções na comunicação política e retração da esfera pública em arenas discursivas digitais. Rio de Janeiro: Lumen Juris, 2021.

SOUZA, Carlos Affonso; LEMOS, Ronaldo; BOTTINO, Celina (Coord.). *Marco Civil da Internet*: jurisprudência comentada. São Paulo: Revista dos Tribunais, 2017.

TERMOS de uso do *Instagram*. Meta, [s.l.], 2022. Disponível em: <https://pt-br.facebook.com/help/instagram/581066165581870>. Acesso em: 10 fev. 2022.