

**Programa Institucional de
Bolsas de Iniciação**

Científica (PIBIC)

Direito

Comportamento Judicial e o
Facebook Oversight Board: O papel
dos comentários públicos

Ruth Gaudêncio da Silva

Orientador: Ivar Alberto Glasherster Martins
Lange Hartmann

São Paulo

2025

Programa Institucional de Bolsas de Iniciação Científica (PIBIC)

Ruth Gaudêncio da Silva

Direito

Comportamento Judicial e o Facebook Oversight Board: O papel dos comentários públicos

Relatório final apresentado ao Programa
Institucional de Bolsas de Iniciação Científica
(PIBIC)

Orientador: Prof. Ivar Alberto Glasherster
Martins Lange Hartmann

São Paulo

2025

Resumo

Esta pesquisa investiga a influência dos comentários públicos de todas as decisões do *Oversight Board* do Facebook até o primeiro quadrimestre de 2025, buscando entender como a quantidade, a origem e o posicionamento desses comentários se relacionam com o resultado dos casos analisados. O objetivo é determinar se a participação do público, expressa por meio de comentários, converge ou diverge das decisões finais do conselho. A metodologia empregada foi uma abordagem mista, combinando análise manual inicial para definir os parâmetros de classificação, e posteriormente, com o uso de técnicas de Processamento de Linguagem Natural (PLN) e aprendizado de máquina, com o uso de *Large Language Models (LLMs)* implementados em *Python* e da ferramenta ChatGPT. A análise automatizada visa classificar e analisar o conteúdo dos comentários, permitindo a identificação de padrões nos comentários. Os resultados indicaram taxa de concordância entre a classificação automatizada e a manual, de 90%, demonstrando o potencial da utilização de LLMs para compreender, posteriormente, a influência da participação pública nas decisões do *Oversight Board*. A análise revelou que a participação pública é altamente concentrada em poucos casos de grande repercussão, enquanto a maioria recebe baixo engajamento, o que limita a representatividade da sociedade civil no processo. Além disso, observou-se a sub-representação do Sul Global, evidenciando desigualdades regionais na participação. Por fim, verificou-se que a correlação entre o posicionamento dos comentários públicos e as decisões finais do conselho é limitada, o que sugere relativa autonomia do *Oversight Board* em suas deliberações. Desse modo, estudo contribui para a compreensão do funcionamento e da eficácia do conselho como um mecanismo de governança, com autonomia e legitimidade das suas decisões.

Palavras-chave: Plataformas digitais; Facebook Oversight Board; Decisão Judicial.

Abstract

This paper investigates the influence of public comments on all decisions of Facebook's Oversight Board up to the first quarter of 2025, seeking to understand how the quantity, origin, and positioning of these comments relate to the outcomes of the analyzed cases. The objective is to determine whether public participation, expressed through comments, converges with or diverges from the Board's final decisions. The methodology employed was a mixed approach, combining an initial manual analysis to define classification parameters and, subsequently, the use of Natural Language Processing (NLP) and machine learning techniques, through the application of Large Language Models (LLMs) implemented in Python and the ChatGPT tool. The automated analysis aimed to classify and examine the content of the comments, enabling the identification of patterns. The results indicated a 90% agreement rate between the automated and manual classifications, demonstrating the potential of LLMs to further explore the influence of public participation on Oversight Board decisions. The analysis revealed that public engagement is highly concentrated in a few high-profile cases, while the majority receive low levels of participation, limiting the representativeness of civil society in the process. Moreover, the underrepresentation of the Global South was observed, highlighting regional inequalities in participation. Finally, the correlation between the stance of public comments and the Board's final decisions was found to be limited, suggesting a relative autonomy of the Oversight Board in its deliberations. Thus, this study contributes to the understanding of the functioning and effectiveness of the Board as a governance mechanism, with autonomy and legitimacy in its decisions.

Keyword: Digital platforms; Facebook Oversight Board; Court Decision.

Sumário

1. INTRODUÇÃO.....	11
2. METODOLOGIA.....	14
2.1 Introdução ao Oversight Board.....	14
2.2 Análise e classificação dos comentários públicos.....	14
2.3 Discussão da Literatura.....	16
2.4 Análise das decisões.....	18
3. REVISÃO DA LITERATURA.....	20
3.1 Expertise and Participation in the Facebook Oversight Board: From Reason to Will 20	
3.2 Facebook's faces.....	21
4. RESULTADOS DA ANÁLISE DAS DECISÕES.....	23
4.1 Análises descritivas.....	23
4.1.1. Frequência de casos com comentários públicos.....	23
4.1.2. Número de decisões que receberam comentários públicos por ano.....	23
4.1.3. Temas das decisões que receberam comentários públicos por ano.....	24
4.1.4 Distribuição de casos por faixa de número de comentários.....	26
4.1.5 Distribuição de casos por geografia.....	27
4.1.6 Temas e casos com mais comentários públicos.....	28
4.1.7. Temas mais frequentes entre as decisões com comentários públicos.....	30

4.2	Efeitos dos comentários públicos.....	32
4.2.1.	Convergência entre a maioria dos comentários públicos e a decisão do <i>Oversight Board</i>	32
4.2.2.	Relação entre o posicionamento dos comentários e a decisão do <i>Oversight Board</i> .	33
4.2.3.	Proporção entre decisões que receberam comentários públicos e decisões que mencionaram comentários públicos.....	35
4.2.4.	Temas mais recorrentes entre as decisões que mencionam comentários públicos ..	36
5.	CONSIDERAÇÕES FINAIS.....	37
	Referências	39
	Apêndice A – Código de extração dos comentários.....	40
	Apêndice B – Prompt para o LLM com o ChatGPT (modelo GPT-4).....	44
	Apêndice C – Ficha de Leitura: DVOSKIN, Brenda. Expertise and Participation in the Facebook Oversight Board: From Reason to Will.....	46
	Apêndice D – Ficha de Leitura: ARUN, Chinmayi. Facebook’s	50
	Apêndice F – Código de análise dos resultados	54

1. INTRODUÇÃO

O *Oversight Board*, ou Conselho de Supervisão, é um órgão independente criado pela Meta (anteriormente Facebook) que atua desde outubro de 2020. Sua função é decidir sobre a manutenção ou remoção de conteúdos nas plataformas Facebook, Instagram e Threads.

Os casos são selecionados pelos membros do Conselho após os usuários esgotarem o processo de apelação da Meta, seja de conteúdo próprio ou que tenha denunciado. A própria empresa também encaminha casos para análise. O *Oversight Board* visa proteger a liberdade de expressão e outros direitos humanos dos usuários. Suas decisões são vinculativas para a Meta, desde que não violem leis locais, e podem incluir recomendações para reformulação das políticas da empresa, que afetam todos os usuários. O conselho é financiado por um fundo independente, estabelecido com uma doação irrevogável da Meta, e é composto por especialistas e líderes de diversas origens.

Uma vez que um caso é selecionado para análise, inicia-se um processo deliberativo estruturado. Primeiramente, o conselho publica uma descrição do caso em seu site, neste momento, abre-se um período para comentários públicos, geralmente de 14 dias, durante o qual qualquer pessoa interessada pode submeter suas considerações sobre o mesmo. Esta etapa visa adicionar à análise perspectivas diversas da sociedade civil, especialistas e partes interessadas.

Após o encerramento do período de comentários, um painel de cinco membros do conselho é designado para analisar o caso. Eles examinam o conteúdo em questão, as políticas relevantes da Meta, os comentários públicos recebidos e quaisquer informações adicionais fornecidas pela empresa, contratadas ou pelo usuário envolvido. O painel então delibera e chega a uma decisão por maioria. Esta decisão é revisada pelo conselho pleno para garantir consistência com casos anteriores e, uma vez finalizada, é publicada no site do *Oversight Board*, juntamente com uma explicação detalhada do raciocínio por trás da decisão. Todo este processo deve ser concluído dentro de 90 dias após da seleção do caso, embora o conselho possa

estender este prazo em situações excepcionais ou de problemas técnicos (*Oversight Board Bylaws*, 2024).

Portanto, os comentários públicos desempenham um papel importante no processo decisório do *Oversight Board*, para a garantia da legitimidade e a transparência de suas decisões. Este mecanismo permite que um maior número de usuários seja ouvido, incluindo aquelas de grupos minoritários e representantes de diversas culturas, que podem não ter sido adequadamente consideradas nas decisões iniciais de moderação de conteúdo.

Ou seja, ao incorporar essas perspectivas variadas, o conselho enriquece sua análise, potencialmente identificando nuances culturais, contextos específicos ou implicações não evidentes à primeira vista. A abertura para comentários públicos também promove um senso de participação da comunidade global de usuários das plataformas da Meta, contribuindo para decisões supostamente mais representativas.

O objetivo do projeto de pesquisa foi realizar uma análise da relação entre os comentários públicos e as decisões finais do *Oversight Board*. Para isso, buscou-se compreender o grau de convergência ou divergência entre os comentários submetidos e a decisão final emitida pelo conselho em cada caso. E se existe uma correlação perceptível entre a quantidade de comentários favoráveis ou contrários a uma determinada posição e a decisão final do *Oversight Board*. Assim, pretende-se verificar se um volume maior de comentários com um posicionamento específico, exerce influência sobre o resultado da decisão, ou se o conselho mantinha sua independência independentemente da proporção de opiniões recebidas.

A metodologia aplicada na pesquisa envolveu a extração de todos os comentários disponíveis em cada caso julgado pelo *Oversight Board*, abrangendo o período de outubro 2020 a abril de 2025. Inicialmente, duas pesquisadoras realizaram uma classificação manual para estabelecer parâmetros de categorização e determinar uma taxa aceitável de erro. Com base nessa análise preliminar, selecionou-se uma amostra aleatória de 41 comentários para processamento pelo ChatGPT (modelo GPT- 4), visando comparar a equivalência das respostas. Foi obtido um resultado positivo de 90% de concordância, e então procedeu-se à análise automatizada dos casos remanescentes utilizando o mesmo modelo de *LLM*.

Ademais, foi feito um estudo da literatura sobre o tema, por meio de dois artigos. O primeiro artigo, de Dvoskin (2023), explora duas estratégias potenciais para o *Oversight Board* construir sua legitimidade. Onde a autora argumenta que os membros do Conselho atuam como tomadores de decisão apolíticos, apresentando o direito internacional dos direitos humanos como um conjunto de princípios universais que podem ser traduzidos e implementados nas redes sociais. Como alternativa, a autora sugere que o este poderia buscar legitimidade promovendo um maior envolvimento da sociedade civil na elaboração das regras de moderação de conteúdo do Facebook.

O segundo artigo, de Arun (2021), examina a complexidade das relações do Facebook com estados, públicos e sua própria equipe. O autor argumenta que o Facebook se engaja em múltiplas conversas regulatórias paralelas complicadas, pelo fato de que a própria empresa não tem uma única estrutura para decidir sobre o comportamento dos usuários ou na hora de análise de casos como nas políticas de publicação. Além disso, ele destaca que o Facebook tem "muitas faces" - diferentes equipes trabalhando com objetivos distintos e interagindo com diversos ministérios, instituições e organizações da sociedade civil.

Desse modo, o presente trabalho de pesquisa está estruturado da seguinte forma. A primeira seção apresenta a introdução, contextualizando o *Oversight Board* e seu papel na moderação de conteúdo das plataformas da Meta. A segunda seção detalha a metodologia empregada, descrevendo o processo de coleta e análise dos comentários públicos submetidos aos casos do *Oversight Board*. A terceira seção aborda a revisão da literatura, analisando os estudos sobre estratégias de legitimação do Conselho e as relações do Facebook com diferentes atores. A quarta seção expõe os resultados da análise das decisões, com foco na relação entre os comentários públicos e as decisões finais do Conselho, bem como a possível influência do volume de comentários sobre decisão final. A última seção apresenta as conclusões.

2. METODOLOGIA

O trabalho foi estruturado em quatro etapas principais: (i) introdução ao *Oversight Board*, (ii) análise e classificação dos comentários públicos, (iii) discussão da literatura e (iv) análise das decisões. Esta estrutura foi escolhida para proporcionar uma compreensão de como o conselho se apresenta para a sociedade e de como ele atua na prática.

2.1 Introdução ao Oversight Board

Na etapa de introdução ao *Oversight Board*, foram realizados dois encontros com o professor orientador para discutir a criação, estrutura e atuação do conselho. Estas discussões focaram na governança do conselho, no perfil dos membros e nos mecanismos de tomada de decisão. A principal fonte para estas análises foram os materiais disponibilizados no site oficial do conselho, que incluem documentos sobre a estrutura organizacional, biografias dos membros e relatórios detalhados de todos os casos decididos. Além disso, foram discutidas as principais perguntas que poderiam ser respondidas ao final da pesquisa: Como o Conselho pode ampliar sua competência, por meio de interpretação das regras básicas que lhe foram impostas, de maneira gradual e segura, de forma que após 10 ou 20 anos esteja em posição de decidir sobre tipos de casos ou temas que em seu primeiro ano não conseguia decidir? É possível falar em legitimidade (democrática) perante usuários de um órgão privado criado por uma empresa com fins lucrativos?

2.2 Análise e classificação dos comentários públicos

A segunda etapa da metodologia consistiu na análise e classificação dos comentários públicos recebidos para cada caso julgado pelo *Oversight Board*. Com o objetivo de construir uma base de dados, para posteriormente compreender se e como o engajamento público influenciam as decisões em relação ao conteúdo publicado.

Esse processo, envolveu duas abordagens com a linguagem de programação *Python*: *web scraping* automatizado para extrair informações online e download manual dos arquivos PDF contendo os comentários (Apêndice A).

Para isso foi feita a conversão de arquivos PDF em imagens de alta resolução (300 DPI) usando a biblioteca do Pandas *pdf2image*. Em seguida, aplicou-se a tecnologia de Reconhecimento Óptico de Caracteres (OCR) através da biblioteca *pytesseract* para extrair o texto dessas imagens. O *script* foi projetado para identificar e extrair especificamente os resumos curtos e os comentários completos de cada submissão, utilizando expressões regulares para localizar padrões de texto como “*Short summary provided by the commenter*” e “*Full Comment*”. Uma função de limpeza de texto foi implementada para remover informações irrelevantes, como cabeçalhos repetitivos e formatações inconsistentes, assegurando a qualidade e consistência dos dados extraídos. Após a extração e limpeza, os dados foram organizados em um *DataFrame* do *Pandas*, incluindo o nome do arquivo de origem, o resumo do comentário e o comentário completo e seu tamanho, considerando o número de caracteres, para uma classificação manual prévia ao treinamento do modelo.

A etapa seguinte da pesquisa envolveu a classificação manual dos comentários, a fim de determinar qual a taxa de erro aceitável em uma análise manual e qual a melhor maneira de classificar os comentários usando LLM. Para isso, foi selecionada uma amostra diversificada de 15 comentários, distribuídos em tamanhos pequeno, médio e grande conforme a quantidade de caracteres, e abrangendo diferentes casos decididos pelo *Oversight Board*. Nesta primeira fase, utilizou-se um formulário estruturado para avaliar a posição do comentarista em relação à decisão da plataforma e ao conteúdo em questão. O formulário continha os campos: "Número do caso" e questões sobre o posicionamento do comentarista em relação a decisão da plataforma: “*Supports the platform's decision*”, “*Disagrees with the platform's decision*”, e “*The commentator's position is nuclear*”. E a segunda pergunta era o posicionamento do comentarista em relação ao conteúdo publicado: “*Supports removing the content*”, “*Supports keeping the content on the platform*” e “*Unclear what the commentator's position is on the content.*”

Após uma análise inicial, identificou-se que a avaliação do alinhamento do comentarista com a decisão da plataforma introduzia redundância no processo. Assim, a classificação foi refinada, focando exclusivamente na questão “*Commentator's position regarding the content*”, com as opções “*Supports removing the content*”, “*Supports keeping the content on the platform*”, e “*Unclear what the*

commentator's position is on the content". Essa simplificação teve por objetivo concentrar a análise na avaliação da postura do comentarista em relação ao conteúdo em si, otimizando a eficiência e a relevância da classificação para as etapas subsequentes da pesquisa.

Com os parâmetros de classificação definidos, realizou-se uma nova etapa de classificação manual, onde 41 casos foram selecionados aleatoriamente e avaliados independentemente por duas pesquisadoras. Adicionalmente, esses mesmos 41 casos foram submetidos à análise automatizada utilizando o ChatGPT (modelo GPT-4), configurado com um modelo de LLM, com o objetivo de comparar a equivalência das classificações.

O prompt (Apêndice B) instruiu o ChatGPT a resumir o ponto principal de cada comentário, identificar frases-chave e sentimentos que indicassem a posição do comentarista em relação à decisão da plataforma, e classificar o comentário em uma das seguintes categorias: "*Supports removing the content*", "*Supports keeping the content on the platforms*", ou "*Unclear what the commentator's position is on the content*". Além disso, o ChatGPT era solicitado a explicar o raciocínio por trás de cada classificação, referenciando partes específicas do comentário que levaram à sua decisão, e fornecer um nível de confiança (alto, médio ou baixo) para cada classificação. Esse prompt detalhado visava garantir que a análise automatizada replicasse, da forma mais precisa possível, o rigor e a objetividade da análise manual realizada pelas pesquisadoras. Ao final o resultado comparado das classificações foi positivo, com uma taxa de 90% de concordância, sendo, portanto, os mesmos parâmetros utilizados para a classificação de todos os comentários no período analisado.

2.3 Discussão da Literatura

Esta etapa consistiu na leitura, discussão e elaboração de fichas de leitura, contendo: resumo, avaliação crítica, principais citações e referências, de cinco artigos distintos, previamente selecionados pelo professor orientador, conforme demonstrado na Tabela 1.

Tabela 1 - Artigos da discussão da literatura

<p>DESAI, Deven R. KROLL, Joshua A. Trust But Verify: A Guide to Algorithms and the Law. Harvard Journal of Law & Technology. V. 31, n. 1, 2017.</p>
<p>DVOSKIN, Brenda. Expertise and Participation in the Facebook Oversight Board: From Reason to Will. Telecommunications Policy. V. 47, 2023.</p>
<p>O’Kane, Ruby. Meta’s Private Speech Governance and the Role of the Oversight Board: Lessons from the Board’s First Decisions. Stanford Technology Law Review. Vol. 25, 2022.</p>
<p>Tiedeke, Anna Sophia & Fertmann, Martin. A Love Triangle? Mapping Interactions between Internacional Human Rights Institutions, Meta and Its Oversight Board. The European Journal of International Law Vol. 34 no. 4, 2024.</p>
<p>ARUN, Chinmayi. Facebook’s faces. Case Western Reserve Journal of International Law, v. 53, n. 1, p. 1-36, 2021.</p>

Fonte: Elaborado pela autora

Com o objetivo de aprofundar a primeira etapa do trabalho, referente à metodologia de análise dos comentários, a descrição detalhada, juntamente com as fichas de leitura dos artigos de Dvonskin (2023), no Apêndice C, e Arun (2021), no Apêndice D, será apresentada na seção de revisão da literatura. Esses artigos foram selecionados por abordarem diretamente a relevância dos comentários públicos como forma de

manifestação da sociedade civil nas decisões, não apenas do *Oversight Board*, mas também, de forma indireta, nas políticas implementadas pela Meta.

2.4 Análise das decisões

A análise das decisões foi feita em duas etapas, (i) tratamento das bases de dados coletados em Excel e agrupamento de bases usadas durante a coleta e da análise de comentários e (ii) análise dos dados usando a linguagem de programação *Python*, com o objetivo de extrair, organizar e visualizar informações (Apêndice D). A escolha por uma abordagem programática se deu pela necessidade de tratar dados semiestruturados (como listas de países em uma única célula), manipular grandes volumes de texto, e gerar visualizações estatísticas personalizadas com precisão e reprodutibilidade.

Para isso, utilizou-se um conjunto de bibliotecas do Python,: (i) *pandas*, utilizada para leitura, manipulação e transformação dos dados em formato tabular (*DataFrames*), especialmente a partir de arquivos Excel. (ii) *numpy*: para operações vetoriais e de apoio a transformações numéricas. *matplotlib* e *seaborn* (iii): para a construção de gráficos e visualizações customizadas (como gráficos de barras por continente e por país). (iv) *pycountry* e *pycountry_convert*: essenciais para converter nomes de países ou seus códigos ISO em regiões geográficas padronizadas, como continentes. (v) *re* (expressões regulares): utilizada na limpeza de *strings*, separação de países em colunas compostas e na padronização de entradas textuais. (vi) *openpyxl*: para leitura e escrita de arquivos Excel quando necessário, especialmente ao lidar com múltiplas abas da planilha original.

A base de dados foi extraída de um arquivo Excel consolidado contendo informações sobre as decisões do *Oversight Board*. A primeira etapa foi o carregamento da aba principal contendo as decisões, que incluía colunas como *id*, *country*, *decision_title* e *Number of public comments*.

Para análise geográfica, a coluna *country*, que frequentemente incluía múltiplos países por célula, e foram separadas as entradas e remover inconsistências como espaços extras, nomes incompletos ou ausentes. Após a separação, os países foram mapeados para seus respectivos continentes utilizando os pacotes *pycountry* e

pycountry_convert. Isso possibilitou a construção de gráficos de distribuição de decisões por continente, tanto considerando apenas os casos com comentários públicos citados nas decisões quanto o universo completo de 118 decisões analisadas.

Como uma decisão podia abranger mais de um país e, portanto, mais de um continente, foi necessário desdobrar essas associações. Cada ocorrência de país-continente foi tratada como uma linha independente para que a visualização refletisse a pluralidade geográfica dos casos. Essa modelagem permitiu análises mais precisas, como a verificação de que 91 decisões envolviam apenas um continente, enquanto outras envolviam até quatro diferentes.

Também foram gerados filtros adicionais para compreender a relação entre os continentes e o engajamento público, visualizando separadamente os casos que receberam comentários públicos citados nas decisões finais. Esse filtro foi aplicado com base na coluna *Number of public comments*, considerando como válidos os casos com valor maior ou igual a 1. Os dados foram tratados para converter esta coluna para o tipo numérico e preencher valores ausentes com zero.

Todos os gráficos foram construídos com as bibliotecas *matplotlib* e *seaborn*, com ajustes estéticos específicos (remoção de eixos, inserção de rótulos percentuais, cores padronizadas), garantindo uma apresentação clara dos dados e facilitando a leitura e comparação entre categorias.

Por fim, todos os scripts desenvolvidos foram organizados em células bem documentadas, com comentários explicativos que descrevem a lógica por trás de cada etapa. Dessa forma, o código pode ser facilmente reutilizado, adaptado ou replicado em outras análises que envolvam dados semelhantes. O *notebook* para replicação das análises pode ser solicitado à autora.

3. REVISÃO DA LITERATURA

3.1 Expertise and Participation in the Facebook Oversight Board: From Reason to Will

No artigo: “*Expertise and Participation in the Facebook Oversight Board: From Reason to Will*” Brenda Dvoskin (2023) examina criticamente as estratégias de legitimação adotadas pelo *Oversight Board*, argumentando que a ênfase no Direito Internacional dos Direitos Humanos (DIDH) como ferramenta central de tomada de decisão resulta em um processo tecnocrático e restritivo. Embora o uso do DIDH possa conferir uma aparência de objetividade e neutralidade, a autora argumenta que ele desconsidera aspectos culturais e contextuais relevantes para os usuários da plataforma, especialmente aqueles do Sul Global.

Portanto, a autora questiona se o DIDH, originalmente concebido para regular a conduta de governos, é adequado para lidar com os desafios únicos da moderação de conteúdo em plataformas digitais, onde o alcance global e a natureza dinâmica da comunicação exigem uma abordagem mais flexível e adaptada às diversas realidades locais.

Como solução para esse problema, a autora propõe que o Conselho fortaleça o engajamento com a sociedade civil, promovendo uma participação mais ampla e inclusiva nas decisões de moderação de conteúdo. A autora sugere que o conselho atue como uma ponte entre a empresa e a sociedade, garantindo que as vozes de grupos subrepresentados sejam ouvidas e consideradas. Embora reconheça as limitações dessa abordagem, como a dependência de financiamento externo e o risco de manipulação por parte da Meta, ela argumenta que a participação da sociedade civil tem o potencial de redistribuir o poder e tornar as decisões do Oversight Board mais legítimas e representativas dos interesses públicos.

Portanto, compreender a influência real dos comentários públicos – atualmente o principal espaço para a manifestação da sociedade civil em relação às políticas de moderação de conteúdo da Meta – nas decisões do *Oversight Board*, permitirá validar argumentos sobre a eficácia desse canal de participação. A análise busca determinar se a proposta de fortalecer a participação da sociedade civil já é eficaz em sua forma

atual ou se outras soluções são necessárias para torná-la mais representativa e impactante.

3.2 Facebook's faces

Chinmayi Arun (2021) explora no artigo como Facebook mantém relações com diferentes atores, incluindo estados, públicos e seus próprios funcionários. A autora argumenta que o Facebook não é uma entidade monolítica, mas sim uma organização multifacetada com equipes e indivíduos que trabalham em direção a diferentes objetivos e prioridades. Essa perspectiva revela que o Facebook se adapta e responde de maneiras distintas conforme o poder e a influência desses.

O artigo também considera a importância da diversidade de vozes dentro do Facebook e como elas moldam as políticas de moderação de conteúdo da plataforma. Para isso, a autora argumenta que o engajamento do Facebook com os estados vai além do cumprimento das leis locais, envolvendo negociações contínuas sobre uma variedade de questões, como proteção de dados, aprovação regulatória de novos produtos e tributação. Essa perspectiva levanta questões sobre a independência do Oversight Board em relação aos interesses comerciais e políticos do Facebook, e sobre sua capacidade de tomar decisões imparciais que atendam aos interesses da sociedade civil.

Desse modo, ao analisar as dinâmicas de poder em jogo nas relações do Facebook com diferentes atores, se enfatiza a necessidade de compreender como estados e públicos mais influentes têm maior capacidade de moldar as políticas da plataforma, enquanto grupos menos poderosos podem ser marginalizados. Essa perspectiva crítica é relevante para a análise dos comentários públicos nas decisões do Oversight Board, pois levanta questões sobre a representatividade e a equidade do processo de consulta pública, assim como e sobre a capacidade de diferentes grupos da sociedade civil de influenciarem as decisões do conselho. Uma vez que nem os próprios Estados conseguem atuar em relação aos Padrões de Comunidade da Meta e conseqüentemente na moderação de conteúdo.

Nesse contexto, a análise dos comentários públicos se torna uma ferramenta importante para avaliar se o Oversight Board está cumprindo seu papel de garantir

que as vozes de todos os grupos sejam ouvidas e consideradas nas decisões de moderação de conteúdo do Facebook ou se as decisões se aproximam de um padrão para manter a legitimidade da empresa frente ao mercado.

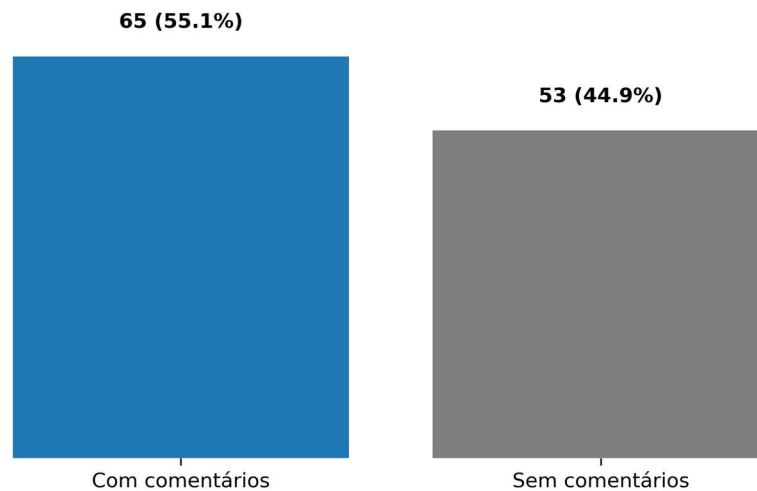
4. RESULTADOS DA ANÁLISE DAS DECISÕES

4.1 Análises descritivas

4.1.1. Frequência de casos com comentários públicos

A primeira análise realizada, teve por objetivo entender o número de decisões que receberam comentários públicos. Esse número é necessário para a análise posterior de qual a relação dos comentários públicos com a decisão e para compreender se os usuários costumam usar o recurso de envio de comentários públicos para o conselho.

Figura 1. Número de decisões que receberam comentários públicos



Fonte: Elaborado pela autora

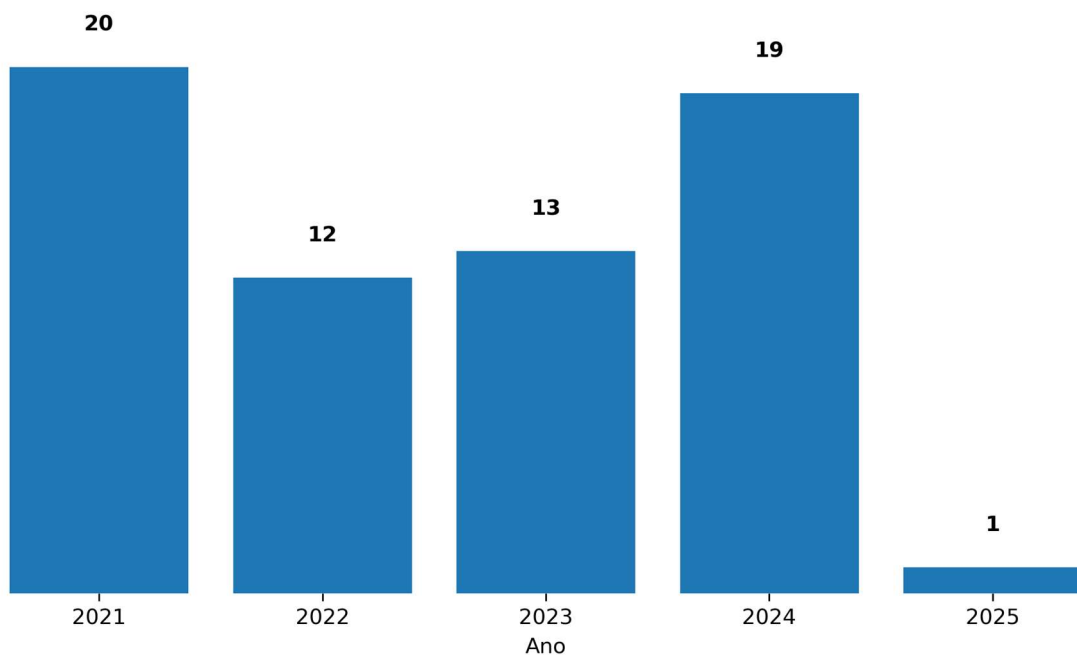
A Figura 1 mostra quantas decisões do *Oversight Board* receberam comentários públicos versus não receberam, no período analisado. Dos 118 casos, 65 (55,1%) foram classificados como: “Com comentários” e 53 (44,9%) como: “Sem comentários”.

4.1.2. Número de decisões que receberam comentários públicos por ano

A Figura 2 demonstra, ano a ano, quantas decisões receberam comentários públicos, ao longo do período analisado. Logo, é possível notar que mesmo que a porcentagem dos casos que recebeu comentários públicos seja pouco maior de 50%, os usuários

utilizam desse recurso desde a instauração do conselho. O gráfico demonstra que nos anos de 2021 e 2024 as decisões receberam mais comentários públicos.

Figura 2. Número de casos que receberam comentários públicos por ano (2021 – 2025)

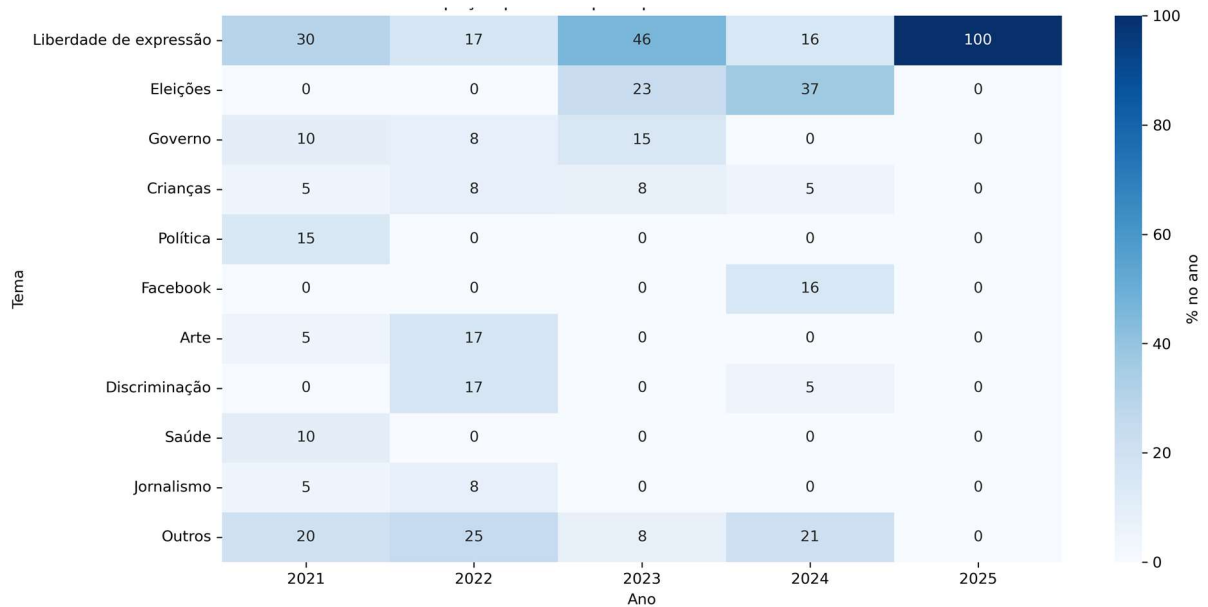


Fonte: Elaborado pela autora

4.1.3. Temas das decisões que receberam comentários públicos por ano

A fim de investigar se havia algum tema responsável pelos anos com maior volume de decisões que receberam comentários, foi construído um *heatmap* de participação por tema principal em cada ano. Para tal, foram consideradas apenas das decisões que receberam ao menos um comentário público; a partir da coluna de temas, foi tomado o primeiro termo listado como “tema principal” do caso; para manter a legibilidade, preservada as dez categorias mais frequentes no período e agregadas as demais em “Outros”; em seguida, calculamos, para cada ano, a participação percentual de cada tema entre as decisões daquele ano e representados esses percentuais no gráfico, conforme a figura 3.

Figura 3. Temas das decisões que receberam comentários públicos por ano



Fonte: Elaborado pela autora

Assim, a leitura da Figura 3 indica que o tema “Liberdade de expressão” é recorrente em todo o período e que “Eleições” tem maior destaque em 2024. Porém, não aparece um único tema que, de forma consistente, explique os anos com maior volume (2021 e 2024).

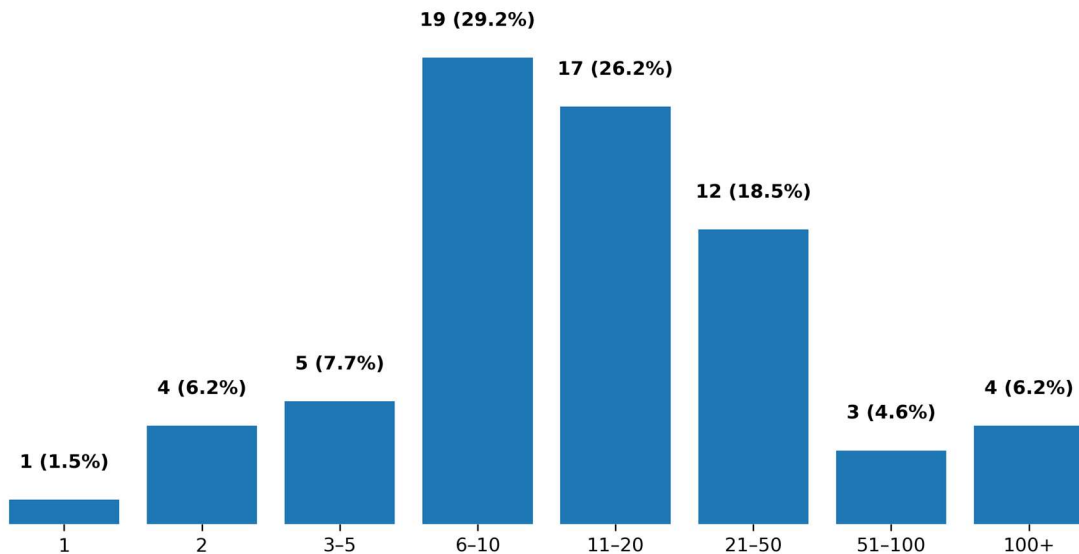
Em 2024, uma hipótese é de o aumento esteja mais ligado à composição — maior proporção de casos eleitorais, enquanto 2021 apresenta uma agenda mais diversificada e pode ter tido influência da pandemia de COVID-19, uma vez que alguns recursos/decisões não foram julgados no período, o que pode ter levado a maior engajamento dos usuários em participar da discussão dos casos a fim de receberem uma conclusão para seus casos.

Ademais, cabe mencionar que essas conclusões são sugestivas, não conclusivas. Foram consideradas na análise: as amostras por ano e por tema, o uso de tema principal desagrupado e considerado os dez temas com maior ocorrência — uma vez que os casos são multitemáticos e a agregação que comprime a cauda longa; além disso, o *heatmap* descreve composição, não quantidade de comentários por caso.

4.1.4 Distribuição de casos por faixa de número de comentários

O gráfico da Figura 4, descreve a distribuição das 65 decisões que receberam comentários públicos (≥ 1), agrupadas por faixas de volume.

Figura 4. Distribuição de casos por faixa de número de comentários



Fonte: Elaborado pela autora

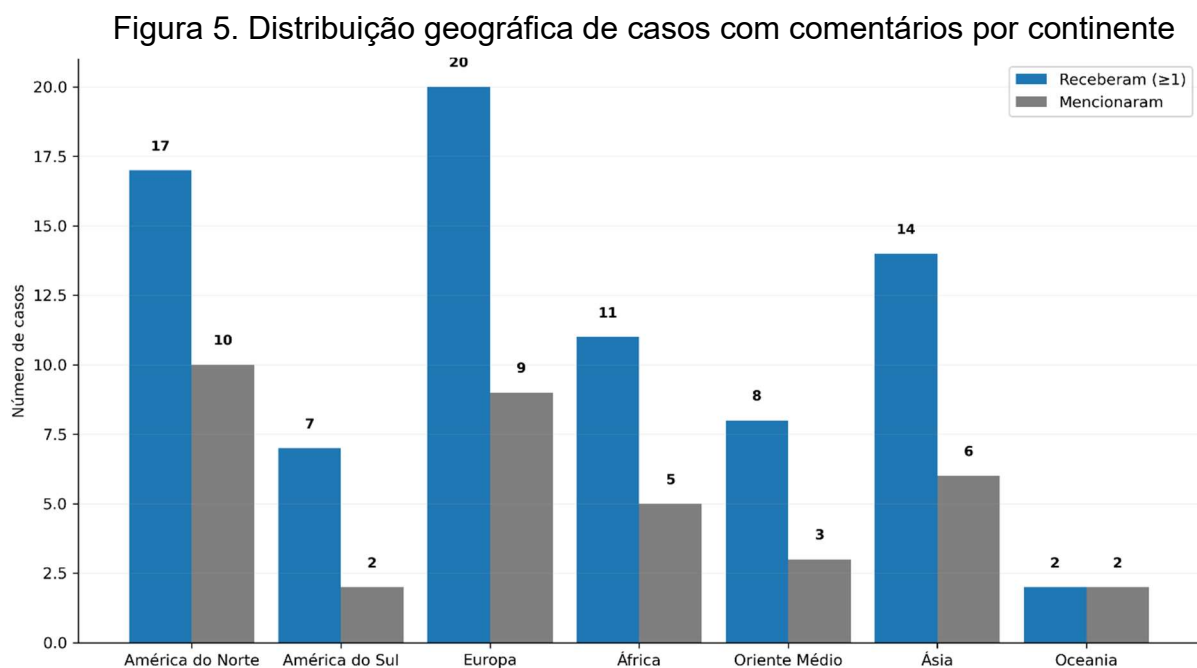
A análise permite observar concentração em níveis moderados: as faixas 6–10 e 11–20 concentram, respectivamente, 19 (29,2%) e 17 (26,2%) casos — juntas, 55,4% do total. A faixa 21–50 ainda é relevante (12 casos; 18,5%). Na outra ponta, há poucos casos com baixa mobilização (1 comentário: 1 caso; 2 comentários: 4 casos; 3–5: 5 casos) e uma cauda longa de alta mobilização (51–100: 3 casos; 100+: 4 casos, somando 10,8%).

Em termos substantivos, isso indica que a participação pública ocorre com frequência, mas tipicamente em patamares médios (entre 6 e 20 comentários por caso). Os poucos casos muito altos formam a cauda e podem ser outliers para a média. Em síntese, o engajamento público é comum, porém os episódios de grande mobilização são exceções, não a regra.

Portanto, a presença de comentários é frequente, porém tipicamente não massiva: a maioria dos casos acumula entre 6 e 20 comentários, e os episódios de grande mobilização, com mais de 100 comentário são exceções.

4.1.5 Distribuição de casos por geografia

A Figura 5, demonstra a distribuição de casos por faixa de número de comentários, considerando os marcadores: “recebeu comentários” (quando *Number of public comments* ≥ 1) e “mencionou comentários” (quando *Is any public comment specifically mentioned?* = Yes). A classificação dos países foi feita com a classificação dos países em que o conteúdo foi postado, na coluna *country* – em que citava mais de um país, dependendo do caso. A fim de se analisar a diferença entre o Norte e o Sul global, além do agrupamento dos países por continente, também foi feita a separação entre explicitamente América do Norte e América do Sul, e entre Ásia de Oriente Médio. Sendo cada decisão contabilizada uma vez por continente citado (ex.: classificação Brasil e EUA, aparece tanto em América do Sul quanto em América do Norte).



Os resultados indicam que a maior concentração de decisões com comentários ocorre na Europa (20 casos) e na América do Norte (17), seguidas por Ásia (14) e África (11); América do Sul (7) e Oceania (2) aparecem com menor volume. Entre quem recebeu comentários, a proporção de decisões que mencionam esses comentários no texto final varia por continente: América do Norte $\approx 58,8\%$, Europa = $45,0\%$, África $\approx 45,5\%$, Ásia $\approx 42,9\%$, Oriente Médio = $37,5\%$, América do Sul $\approx 28,6\%$ e Oceania = 100% .

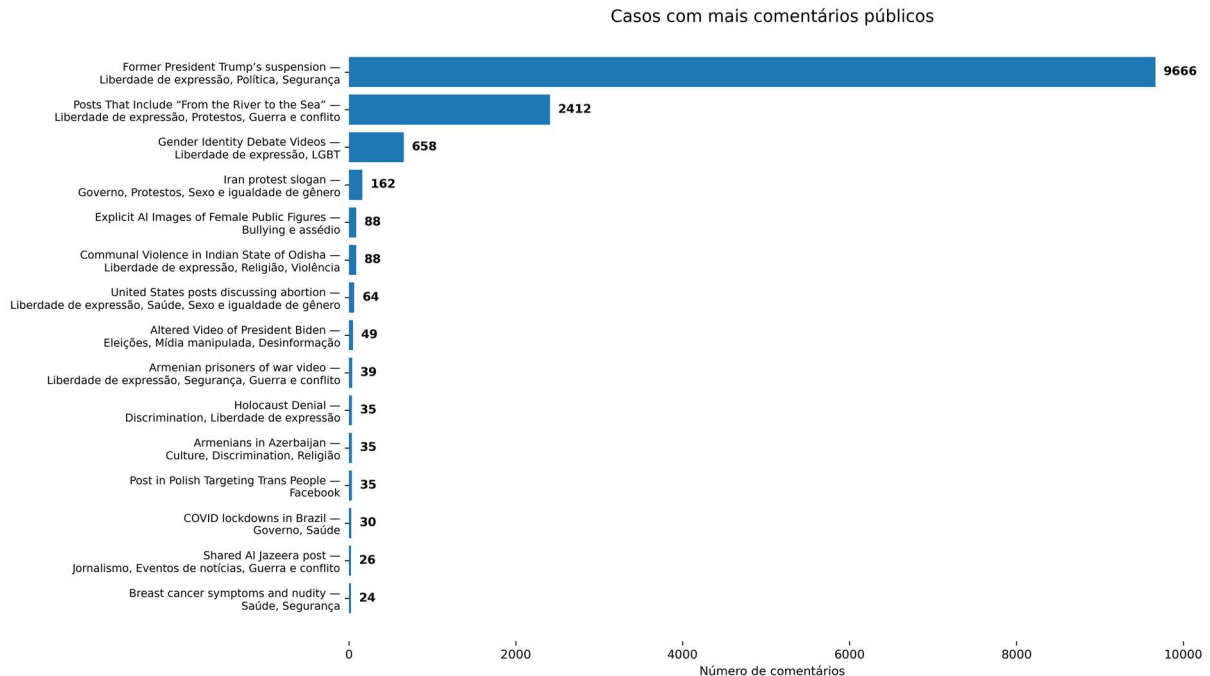
A análise reforçar o argumento de Arun (2021) sobre as relações desiguais do Facebook com diferentes atores, sublinhando que “nem todos têm o mesmo poder de influência” e que estados e públicos mais influentes têm maior capacidade de moldar as políticas da plataforma. A autora questiona se o Board está realmente ouvindo todas as vozes ou se busca “manter a legitimidade da empresa frente ao mercado”.

A concentração dos comentários mencionados em temas como “Eleições”, “Protestos”, “Guerra e conflito” e “Governo” ilustra a tese de Arun. Estes são temas frequentemente associados a debates de alto perfil público e a interesses estatais ou de mercado significativos. Se o Board tende a mencionar comentários em casos que são sensíveis/visíveis ou que exigem justificativa pública mais robusta, isso pode indicar uma preocupação em responder a pressões de atores mais influentes ou em manter a imagem da empresa perante o mercado e os reguladores. Indo de encontro ao que a autora argumenta sobre o Facebook se adaptar e responder de maneiras distintas conforme o poder e a influência desses atores.

4.1.6 Temas e casos com mais comentários públicos

A Figura 6, indica quais os casos que receberam maior volume de comentários públicos no âmbito do *Oversight Board*, evidenciando forte concentração em poucos episódios de alta repercussão (*outliers*).

Figura 6. Distribuição de casos por faixa de número



Fonte: Elaborado pela autora

Da análise do ranking, o caso *"Former President Trump's suspension"* (Liberdade de expressão, Política, Segurança) lidera de forma expressiva, com 9.666 comentários, correspondendo a uma mobilização pública sem precedentes em comparação com os demais. Em segundo lugar aparece *"Posts That Include 'From the River to the Sea'"* (Liberdade de expressão, Protestos, Guerra e conflito), com 2.412 comentários, seguido à distância por *"Gender Identity Debate Videos"* (Liberdade de expressão, LGBT), que registrou 658 comentários.

Após esses três primeiros, observa-se um declínio acentuado no engajamento: *"Iran protest slogan"* (Governo, Protestos, Sexo e igualdade de gênero) reúne 162 comentários, enquanto *"Explicit AI Images of Female Public Figures"* (Bullying e assédio) e *"Communal Violence in Indian State of Odisha"* (Liberdade de expressão, Religião, Violência) aparecem com 88 cada. Em seguida, surgem *"United States posts discussing abortion"* (Liberdade de expressão, Saúde, Sexo e igualdade de gênero) com 64, *"Altered Video of President Biden"* (Eleições, Mídia manipulada,

Desinformação) com 49, e “*Armenian prisoners of war video*” (Liberdade de expressão, Segurança, Guerra e conflito) com 39.

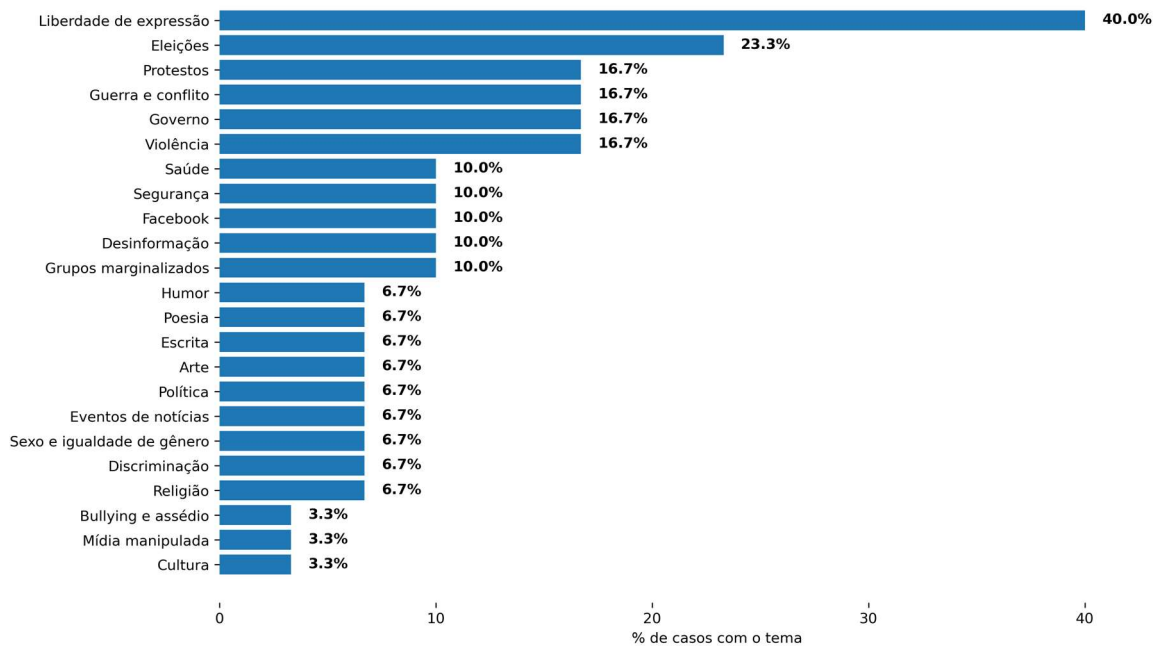
Na faixa dos casos de menor repercussão, encontram-se “*Holocaust Denial*” (Discriminação, Liberdade de expressão), “*Armenians in Azerbaijan*” (Cultura, Discriminação, Religião) e “*Post in Polish Targeting Trans People — Facebook*” (Liberdade de expressão, LGBT), todos com 35 comentários cada. A eles se somam “*COVID lockdowns in Brazil*” (Governo, Saúde) com 30, “*Shared Al Jazeera post*” (Jornalismo, Eventos de notícias, Guerra e conflito) com 26, e por fim “*Breast cancer symptoms and nudity*” (Saúde, Segurança) com 24.

A análise da distribuição temática dos casos que receberam comentários públicos e, mais especificamente, dos que tiveram comentários explicitamente mencionados pelo Oversight Board, complementa de forma significativa a discussão anterior sobre a concentração de volume de comentários e suas implicações para os argumentos de Dvoskin (2023), que critica a ênfase do *Oversight Board* no Direito Internacional dos Direitos Humanos (DIDH), argumentando que essa abordagem pode levar a um processo tecnocrático que desconsidera aspectos culturais e contextuais relevantes para os usuários, especialmente aqueles do Sul Global. Ela defende um maior engajamento da sociedade civil para garantir que as vozes de grupos subrepresentados sejam ouvidas. E a predominância da "Liberdade de expressão" e de temas político-sociais amplos como “Eleições” e “Guerra e conflito”, reforça essa necessidade.

4.1.7. Temas mais frequentes entre as decisões com comentários públicos

A Figura 8 apresenta a proporção de decisões em que cada tema aparece entre as decisões analisadas. Para a análise, cada tema foi contabilizado uma única vez por decisão, logo, um tema pode aparecer em mais de uma decisão, motivo pelo qual os percentuais não somam 100%.

Figura 8. Temas mais frequentes entre as decisões com comentários públicos



Fonte: A autora

Assim, a imagem apresenta o tema Liberdade de expressão como predominante, presente em cerca de 40% dos casos. Em seguida, observa-se um bloco de temas que concentra alta visibilidade pública: Eleições (23,3%), Protestos (16,7%), Guerra e conflito (16,7%), Governo (16,7%) e Violência (16,7%). Enquanto a categoria intermediária é composta por temas como Saúde, Segurança, Desinformação, Grupos marginalizados e Facebook (10% cada), que aparecem com relevância significativa, mas em menor frequência. Finalmente, há um conjunto distribuído de tópicos mais específicos — como Humor, Poesia, Escrita, Arte, Política, Discriminação, Religião, Bullying e assédio, Cultura, Mídia manipulada — cada um com incidência entre 3,3% e 6,7%, sendo, portanto, questões complementares que, embora relevantes, não alcançam a mesma escala de mobilização pública.

Uma hipótese para a predominância dos temas liberdade de expressão, política e segurança nos casos analisados pode ser explicada pelo modo como as plataformas digitais amplificam conteúdos polarizadores e de alto risco social, favorecendo tanto o engajamento orgânico dos cidadãos quanto maior atenção institucional. Como destacam Larreguy e Raffler (2025), a internet e as redes sociais reduziram barreiras de entrada para a circulação de informações e possibilitaram a coordenação de

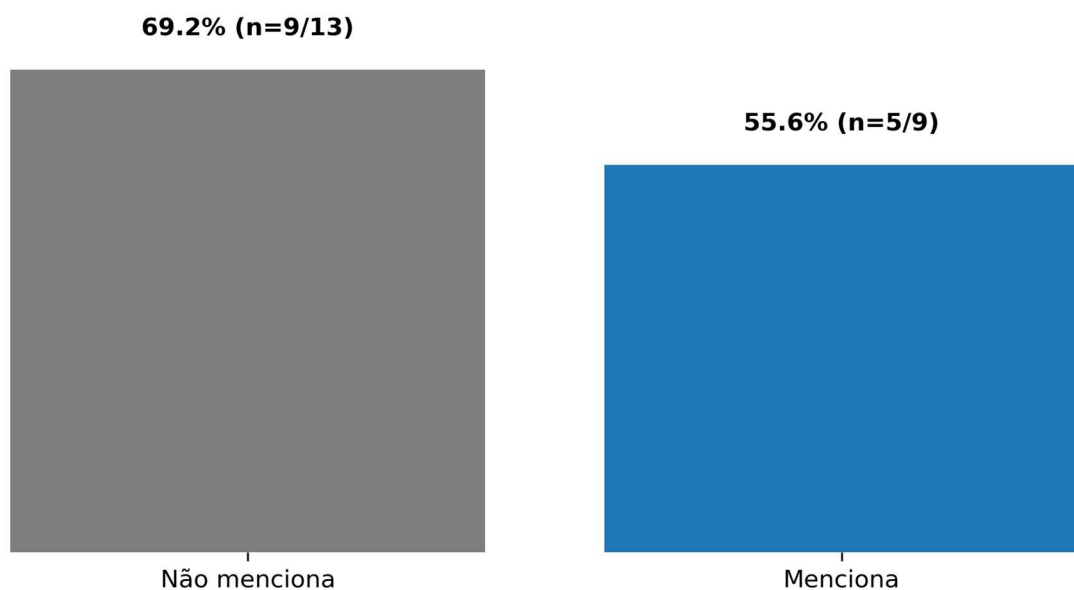
comportamentos em larga escala, sobretudo em contextos de eleições, protestos, guerras e violência, em que a pressão sobre instituições é mais intensa. Em contraste, temas sociais de impacto direto, como saúde ou minorias, mobilizam audiências mais segmentadas, o que explica sua menor representatividade. Assim, os casos de maior visibilidade refletem não apenas disputas de liberdade de expressão, mas também o papel das plataformas como arena central de conflitos políticos e sociais globais, em que decisões de moderação extrapolam o ambiente digital e impactam diretamente a governança democrática.

4.2 Efeitos dos comentários públicos nas decisões

4.2.1. Convergência entre a maioria dos comentários públicos e a decisão do *Oversight Board*

Para a análise de convergência entre a maioria dos comentários públicos recebido sobre o caso e a decisão final do *Oversight Board*, foi considerada a amostra de 22 decisões nas quais os comentários públicos foram classificados, suportando a decisão de manter ou remover o conteúdo em relação à decisão do conselho de manter ou remover o conteúdo. A conclusão foi de que o conselho coincidiu com a maioria dos comentários em 63,6% das decisões e divergiu em 36,4%.

Figura 9. Convergência do Oversight Board com a maioria dos comentários sobre a decisão



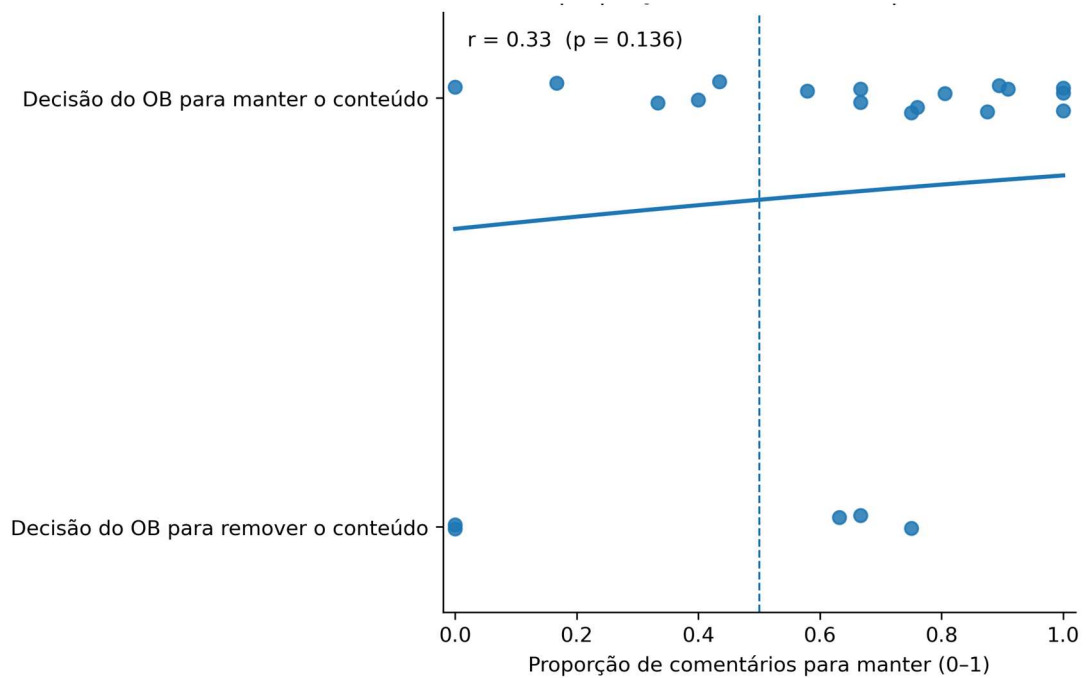
Fonte: Elaborado pela autora

Embora a tabela mostre uma assimetria descritiva (80% de manter quando a maioria foi “manter”, contra 28,6% de remover quando a maioria foi “remover”), os testes de independência não indicam associação no conjunto analisado ($\chi^2 = 0,00$; $p = 1,00$; Fisher OR = 1,60; $p = 1,00$). Isso é consistente com a orientação de Levin, Fox & Forde: o qui-quadrado testa desvio entre observado e esperado sob independência e, com tamanho amostral pequeno e células com baixas frequências esperadas, o teste (ou mesmo o Fisher) tem poder limitado, aumentando a chance de não detectar uma associação real. Além disso, o tamanho de efeito para 2×2 ($\phi = 0,00$) e a concordância além do acaso ($\kappa = 0,093$) sugerem associação/acordo muito fracos. Logo, o N considerado, a evidência é insuficiente para afirmar que a direção da maioria dos comentários esteja associada à decisão do Oversight Board.

4.2.2. Relação entre o posicionamento dos comentários e a decisão do Oversight Board

A Figura 10, apresenta um gráfico de dispersão, e relaciona a proporção de comentários por decisão, que suporta a decisão de manter o conteúdo da publicação (eixo X, 0-1) com a decisão do Oversight Board (eixo Y: manter/remover o conteúdo). A linha pontilhada em 0,5 marca a maioria simples dos comentários e a linha contínua indica a tendência linear. Observa-se uma associação positiva fraca: $r = 0,33$ ($\approx 11\%$ da variância; $r^2 \approx 0,11$), $p = 0,136$. A análise foi feita considerando a amostra de 22 decisões, que possuem comentários válidos classificados (suportam manter a publicação ou suportam remover a publicação), os demais comentários que não se deixavam o posicionamento bem definido, não foram considerados.

Figura 10. Relação entre o posicionamento dos comentários e a decisão do *Oversight Board*



Fonte: Elaborado pela autora

O que indica que casos com maior fração de comentários que suportam manter a publicação, tendem a resultar na decisão do conselho de manter. Os pontos à esquerda de 0,5 (maioria que suporta remover a publicação) em que o *Oversight Board* decide manter, e alguns à direita de 0,5 em que decide remover, sugerindo que o conselho não segue mecanicamente a maioria dos comentários.

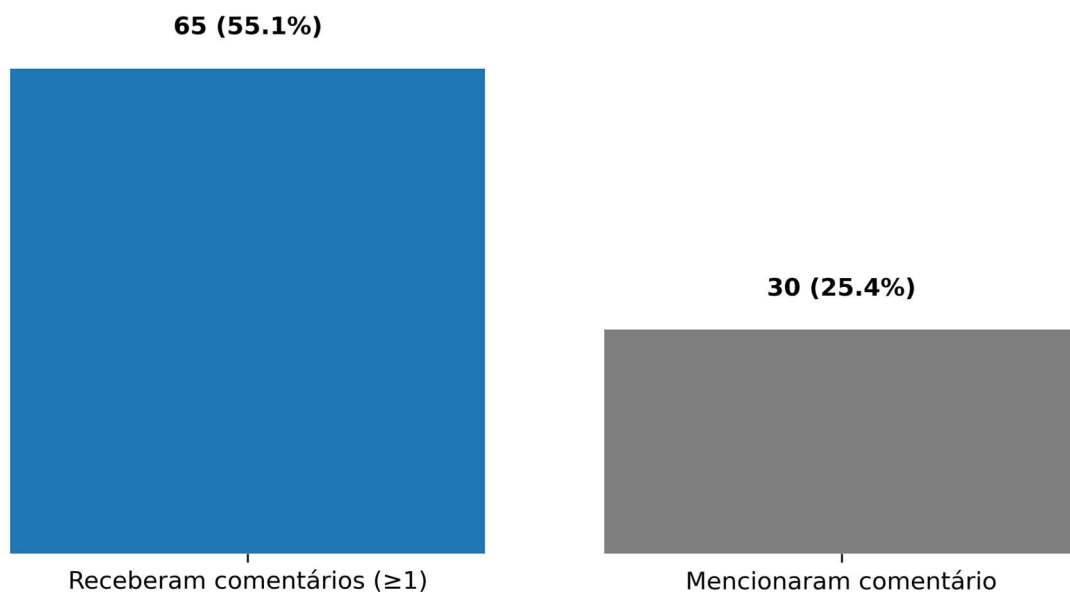
Ademais, a interpretação deve considerar o tamanho amostral reduzido, que resulta em baixo poder estatístico para detectar efeitos moderados. Estimativas simples mostram que, com $N \approx 22$, a chance de detectar um efeito desse tamanho é baixa (~27–32%) como escrevem Levin, Fox & Forde (2012) ao discutir que amostras pequenas aumentam a probabilidade de erro do tipo II (baixo poder) e que aumentar o N é a forma mais direta de melhorar essa detecção.

4.2.3. Proporção entre decisões que receberam comentários públicos e decisões que mencionaram comentários públicos

A fim de compreender melhor como os membros do conselho se relacionam com os comentários públicos recebidos, foi analisada a porcentagem de comentários públicos que foram mencionados em alguma decisão.

Do universo de 118 decisões, 65 receberam ao menos um comentário público (55,1%). Entretanto, apenas 30 decisões mencionaram explicitamente comentários em seus textos (25,4% do total). Olhando só para quem recebeu comentários (n=65), 30 os mencionaram — isto é, 46,2% citaram comentários e 53,8% não citaram.

Figura 11. Proporção entre decisões que receberam comentários públicos e que mencionaram comentários públicos



Fonte: Elaborado pela autora

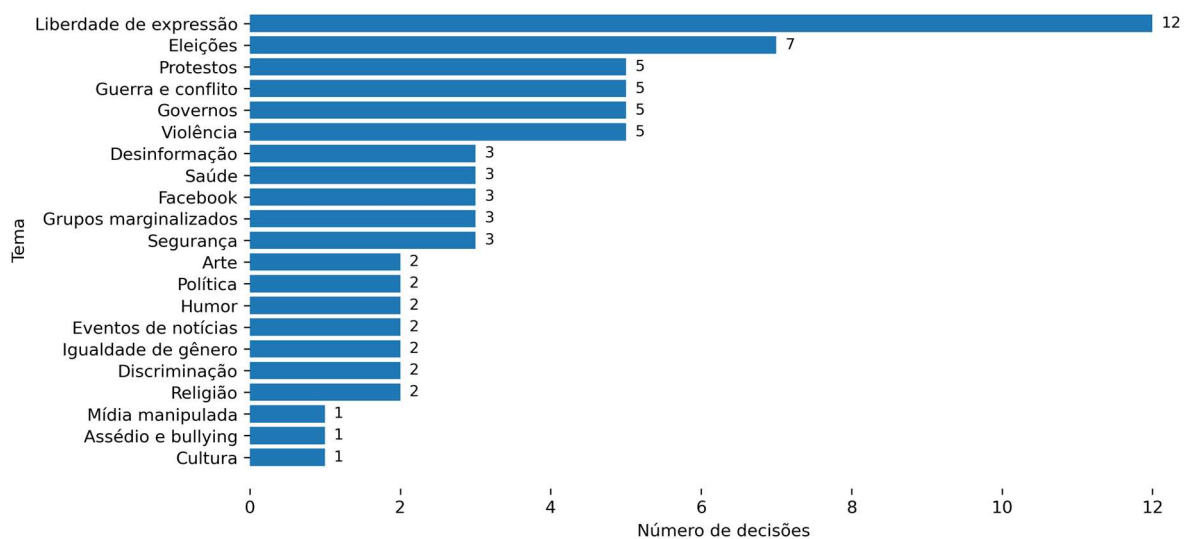
Essa análise sugere que a menção é seletiva, e uma hipótese para tal é que os comentários sejam concentrados em temas específicos que tendem a ter mais repercussão popular.

4.2.4. Temas mais recorrentes entre as decisões que mencionam comentários públicos

De modo a compreender quais os principais temas que receberam comentários (Figura 10), foram analisados os temas mais recorrentes entre as decisões que citaram comentários públicos.

Sendo assim, a figura 12 apresenta, individualmente (se uma decisão tem “Liberdade de expressão, Eleições”, ela entra em ambos) os temas que mais mencionam comentários públicos. Nesse recorte, destacam-se Liberdade de expressão como tema mais recorrente (12 ocorrências), seguido de Eleições (7). Um segundo bloco reúne Guerra e conflito, Violência, Governos e Protestos (cada um com 5). Em patamar menor aparecem Grupos marginalizados, Facebook, Saúde, Segurança e Desinformação (todos com 3), e Política (2).

Figura 12. Temas mais recorrentes entre as decisões que mencionam comentários públicos.



Fonte: A autora

Portanto, a menção a comentários públicos concentra-se em temas fortemente ligados ao debate público (liberdade de expressão e eleições) e em contextos de risco e disputa (conflitos, violência, governos, protestos). Seguindo a mesma tendência da figura 7.

5. CONSIDERAÇÕES FINAIS

A análise dos resultados permite observar que os comentários públicos desempenham um papel limitado, mas não irrelevante, nas decisões do Oversight Board. Os resultados estatísticos demonstram que a direção majoritária dos comentários não determina de forma significativa os resultados finais, uma vez que a concordância além do acaso é mínima, e os testes não identificaram associações estatisticamente relevantes.

Ainda assim, existe uma tendência de que casos com maior apoio a manter conteúdos resultem em decisões nessa direção, o que sugere um efeito indireto, embora não determinante. Essa constatação reforça a ideia de que a participação pública não funciona como um mecanismo de votação direta, mas como uma camada adicional de legitimidade discursiva, utilizada pelo Conselho em situações específicas. A seletividade na mobilização desses comentários, portanto, indica que seu impacto está mais relacionado à função simbólica de reforçar a credibilidade da decisão do que a uma influência substantiva sobre o desfecho.

Contudo, a análise descritiva dos casos em que os comentários foram explicitamente mencionados. A análise da distribuição do volume de comentários públicos, que mostra uma forte concentração em poucos casos de alta repercussão — como a suspensão de Donald Trump e o caso “*From the River to the Sea*” — reforça as críticas de Brenda Dvoskin (2023) e Chinmayi Arun (2021) sobre legitimidade, participação e dinâmicas de poder no *Oversight Board*.

Em vez de uma participação ampla e representativa, observa-se que o engajamento público é altamente assimétrico: enquanto temas de grande visibilidade política global recebem milhares de contribuições, outros casos de igual ou maior relevância para comunidades do Sul Global permanecem marginalizados. Exemplos dessa desigualdade são o contraste entre o engajamento massivo nos casos norte-americanos e a baixa participação em temas como lockdowns de COVID-19 no Brasil ou episódios de violência comunitária em Odisha, na Índia. Nesse sentido, o volume de participação não é apenas um reflexo do interesse dos usuários, mas também das estruturas de poder que determinam quais temas têm legitimidade pública global. O

resultado é que questões cruciais para o Sul Global, como guerras civis, repressões estatais ou restrições de direitos, embora mencionadas nos comentários, não encontram no Board um espaço efetivo de defesa ou reconhecimento.

De forma complementar, Chinmayi Arun (2021) destaca que o Facebook e suas instâncias de governança mantêm relações desiguais com diferentes atores, de modo que “estados e públicos mais influentes têm maior capacidade de moldar as políticas da plataforma, enquanto grupos menos poderosos podem ser marginalizados”. A disparidade no volume de comentários nos casos do Board é uma prova empírica dessa dinâmica: a suspensão de Donald Trump, por exemplo, mobilizou o maior volume de participação pública. Isso sugere que o Board, ao reconhecer publicamente os casos de maior repercussão, pode estar mais voltado a preservar a legitimidade da empresa perante públicos poderosos e reguladores do Norte Global do que a funcionar como uma corte global inclusiva.

Portanto, a distribuição altamente concentrada do engajamento público evidencia um paradoxo geopolítico: os temas que mais impactam o Sul Global — guerras, protestos, repressões governamentais — aparecem entre os mais mencionados, mas esse reconhecimento não se traduz em uma defesa substantiva dos direitos das populações diretamente afetadas. Ao contrário, alinha-se a crítica de que o *Oversight Board* opera sob dinâmicas de poder que privilegiam atores e mercados centrais, sob um discurso de equidade.

Referências

ARUN, Chinmayi. Facebook's faces. *Case Western Reserve Journal of International Law*, v. 53, n. 1, p. 1-36, 2021.

DVOSKIN, Brenda. Expertise and Participation in the Facebook Oversight Board: From Reason to Will. *Telecommunications Policy*, v. 47, 2023.

LARREGUY, Horacio; RAFFLER, Pia J. Accountability in Developing Democracies: The Impact of the Internet, Social Media, and Polarization. *Annual Review of Political Science*, v. 28, p. 413–434, 2025. DOI: <https://doi.org/10.1146/annurev-polisci-033123-015559>.

LEVIN, Jack; FOX, James Alan; FORDE, David R. *Estatística para ciências humanas*. 11. ed. Tradução de Jorge Ritter; revisão técnica de Fernanda Bonafini. Cap. 7 e 10. São Paulo: Pearson Education do Brasil, 2012.

OVERSIGHT BOARD. Oversight Board Bylaws. Disponível em: <https://www.oversightboard.com/wp-content/uploads/2024/03/Oversight-Board-Bylaws.pdf>. Acesso em: 22 fev. 2024.

OVERSIGHT BOARD. São Paulo, Brasil. Disponível em: <https://www.oversightboard.com/>. Acesso em: 22 fev. 2024.

Apêndice A – Código de extração dos comentários

```

import os

import pdf2image

import pytesseract

import re

import pandas as pd

# Define o diretório do script e o diretório dos PDFs a serem processados

script_dir = os.path.dirname(__file__)

pdf_dir = os.path.join(script_dir, 'pdfs')

# Listas para armazenar os dados de todos os arquivos, resumos e comentários

all_files = []

all_summaries = []

all_comments = []

# Função para extrair o conteúdo entre dois padrões especificados

# Recebe como parâmetro o texto completo, o padrão de início e o padrão de fim

# Retorna todos os textos que estão entre esses padrões

def extrair_conteudo(texto, inicio, fim):

    padrao = f'{inicio}(.*?){fim}' # Define um padrão usando expressão regular para capturar o
    conteúdo entre o 'inicio' e o 'fim'

    return re.findall(padrao, texto, re.DOTALL) # Retorna todas as ocorrências encontradas

# Função para limpar o texto extraído de caracteres indesejáveis e seções irrelevantes

# Aplica expressões regulares para fazer substituições e remoções

def limpar_texto(texto):

    # Remove linhas relacionadas ao "Public Comment Appendix" que são irrelevantes

    texto_limpo =
re.sub(r'\n\n(?:Public\s*\Comment\s*\Appendix|PublicCommentAppendix)\s*\|\s*\d+\s*\n', '\n',
texto)

    # Reduz várias linhas em branco seguidas para duas linhas apenas, melhorando a
    formatação do texto

```

```

texto_limpo = re.sub(r'\n{3,}', '\n\n', texto_limpo)

# Remove partes específicas do texto que estão antes da resposta de uma organização
padrao = r'.*?Response on behalf of organization\s*\n*'
texto_limpo = re.sub(padrao, "", texto_limpo, flags=re.DOTALL)

return texto_limpo.strip() # Remove espaços extras no início e fim do texto

# Função principal do script

def main():

    # Loop através de todos os arquivos no diretório 'pdfs'

    for file_name in os.listdir(pdf_dir):

        # Processa apenas arquivos que possuem extensão PDF

        if file_name.lower().endswith('.pdf'):

            file_path = os.path.join(pdf_dir, file_name)

            # Converte as páginas do PDF em imagens

            # '300' é a resolução em DPI, que afeta a qualidade do OCR

            pages = pdf2image.convert_from_path(file_path, 300)

            # Inicializa uma lista para armazenar o texto extraído de cada página

            all_text = []

            # Itera sobre as páginas convertidas, começando pela terceira página (índice
for page_number, page_data in enumerate(pages[2:], start=3):

                # Salva a imagem da página com um nome baseado no número da página

                page_image_path = f'page_{page_number}.png'

                page_data.save(page_image_path, 'PNG')

                # Faz OCR na imagem da página usando pytesseract para extrair o texto

                text = pytesseract.image_to_string(page_data)

                all_text.append(text)

                # Remove a imagem após a leitura para economizar espaço

```

```

os.remove(page_image_path)

# Junta o texto de todas as páginas em uma única string

texto = "\n".join(all_text)

# Extrair os resumos usando o padrão definido de início e fim

resumos = extrair_conteudo(texto, 'Short summary provided by the commenter', 'Full
Comment')

# Extrair os comentários completos usando o padrão definido de início e fim comentarios =
extrair_conteudo(texto, 'Full Comment', 'Link to Attachment')

# Limpar os resumos e comentários extraídos

resumos_limpos = [limpar_texto(resumo) for resumo in resumos]

comentarios_limpos = [limpar_texto(comentario) for comentario in comentarios]

# Adicionar os dados extraídos às listas globais

all_files.extend([file_name] * len(resumos)) # Cria uma entrada de arquivo para cada resumo

all_summaries.extend(resumos_limpos) # Adiciona os resumos limpos

all_comments.extend(comentarios_limpos) # Adiciona os comentários limpos

# Imprime o nome do arquivo para indicar que foi processado

print(f"Processado: {file_name}")

# Criar um DataFrame final com todas as informações extraídas

df = pd.DataFrame({

    'Arquivo': all_files, # Nome do arquivo de origem

    'Resumo': all_summaries, # Resumo do comentário

    'Comentário Completo': all_comments, # Comentário completo

})

# Limpar espaços em branco no início e fim de cada célula

df = df.applymap(lambda x: x.strip() if isinstance(x, str) else x)

# Salvar o DataFrame em um arquivo CSV e também em um arquivo Excel

```

```
df.to_csv('todos_comentarios.csv', index=False)

df.to_excel('todos_comentarios.xlsx', index=False)

# Imprime uma mensagem de conclusão

print("Processamento concluído. Resultados salvos em 'todos_comentarios.csv'")

# Executa a função principal

if __name__ == "__main__":

    main()
```

Apêndice B – Prompt para o LLM com o ChatGPT (modelo GPT-4)

""You are a helpful assistant whose main job is to classify comments related to Facebook or Instagram's content moderation decisions.

I will provide you with one comment at a time. Please follow these steps to analyze and classify each comment:

1. Summarize the main point of the comment in one sentence.
2. Identify any key phrases or sentiments that indicate the commenter's stance on the decision.
3. Consider the context and tone of the comment.
4. Based on your analysis, classify the comment.

You should classify the comment based on the desired action for the content. *Again, use the full description from the options provided below*:

- A. Supports removing the content.
 - B. Supports keeping the content on the platform.
 - C. Unclear what the commentator's position is on the content.
 - D. The comment is in a language that I am unable to understand, even with translation tools.
 - E. The comment mentions that a file was attached, and without being able to review the file, it's unclear what the commentator's position is regarding the content.
5. Explain your reasoning for each classification, referencing specific parts of the comment that led to your decision.
 6. When providing your classification, *do not truncate or alter the category labels*. Ensure that you repeat the entire classification phrase as it appears in the options provided.
 7. If there's any ambiguity, mention alternative interpretations and why you didn't choose them.
 8. Provide a confidence level (high, medium, or low) for each classification.

Please provide your analysis and classification in the following format:

Summary: [One-sentence summary of the comment]

Key phrases/sentiments: [List of relevant phrases or sentiments]

Context and tone: [Brief description of the context and tone]

Classification: [Category name]

Reasoning: [Explanation for the classification]

Confidence: [High/Medium/Low]

Ambiguity (if any): [Discussion of alternative interpretations]

Important notes:

- Strive to be objective and avoid personal biases when analyzing comments.
- If a comment is not in English, please translate it first and then proceed with the analysis.

Here is the public comment for you to classify:\n""

Apêndice C – Ficha de Leitura: DVOSKIN, Brenda. Expertise and Participation in the Facebook Oversight Board: From Reason to Will.

DVOSKIN, Brenda. **Expertise and Participation in the Facebook Oversight Board: From Reason to Will**. Telecommunications Policy. V. 47, 2023. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0308596122001653>

Resumo

O artigo aborda duas principais estratégias para a legitimidade do Facebook Oversight Board (FOB), analisando a abordagem em relação ao Direito Internacional dos Direitos Humanos (DIDH) e a participação pública. Onde o FOB se posiciona como um órgão independente para decisões de moderação de conteúdo, apresentando-se como apolítico onde os membros justificam suas decisões como interpretações técnicas de normas universais, para projetar objetividade e neutralidade. Contudo, o texto argumenta que essa abordagem, ao depender exclusivamente de especialistas em direitos humanos, isola o processo de governança de uma política mais participativa.

Inicialmente, o artigo contextualiza a criação do FOB como uma resposta à desconfiança pública em relação ao poder das plataformas digitais na moderação de conteúdo. Estruturado como um “tribunal”, o FOB busca lidar com disputas relacionadas a remoções de conteúdo e sugerir mudanças nas políticas do Facebook. Assim, a aplicação do DIDH é uma estratégia que reforça a legitimidade do conselho, uma vez que os princípios de direitos humanos possuem ampla aceitação global. No entanto, a autora questiona se esse enquadramento realmente reflete os interesses públicos, dado que o DIDH não resolve totalmente os desafios únicos impostos pelas redes sociais, como o alcance global em que o contexto precisa ser levado em consideração, e a natureza dinâmica da comunicação online que é diferente da proteção editorial de veículos de informação.

Em seguida, o texto propõe uma alternativa: o FOB poderia aumentar sua legitimidade promovendo a participação direta da sociedade civil na formulação das decisões. Essa abordagem permitiria que o conselho se apresentasse como uma ponte entre a empresa e a sociedade, representando múltiplos pontos de vista. Entre as propostas estão a inclusão de organizações de diferentes setores nos debates e a adoção de processos transparentes que mostrem como os insumos externos influenciam as decisões do Facebook. Esse modelo participativo tem o potencial de redistribuir o poder, permitindo que vozes sub-representadas influenciem as regras de moderação de conteúdo.

Contudo, a autora também reconhece limitações significativas dessa abordagem. A dependência de financiamento externo pelas organizações da sociedade civil pode restringir a representatividade e a frequência de participação. Além disso, as consultas públicas podem ser manipuladas pelo Facebook, que mantém controle sobre os temas abordados e os participantes convidados. Ainda assim, o artigo sugere que padrões para consultas e relatórios públicos poderiam aumentar a responsabilidade do Facebook em suas práticas de engajamento com stakeholders.

Como conclusão o texto descreve que, embora o modelo atual baseado no DIDH ofereça vantagens em termos de objetividade, ele não aborda as preocupações mais amplas sobre governança participativa e a influência corporativa. Ao propor a transição para uma abordagem que priorize a participação da sociedade civil, a autora sugere que o FOB poderia não apenas legitimar suas ações, mas também teria potencial para que as decisões não sejam apenas de um pequeno grupo de especialistas como representante de diferentes interesses.

Avaliação Crítica

O artigo fundamenta bem como se garantir legitimidade e também quais as estratégias de legitimidade o FOB tem usado no seu processo de análise de decisões: Direito Internacional dos Direitos Humanos (DIDH) e uma sutil participação da sociedade civil. Sendo assim, um dos aspectos relevantes do artigo é exatamente a crítica feita a ferramenta central de legitimação do conselho ser apenas a do DIDH, uma vez que este foi pensado para governos e não para empresas. Assim, além das decisões tenderem a ser tecnocráticas e restritivas, por não levarem em consideração

aspectos culturais dos usuários da plataforma que tem direitos diferentes e também fazem uso das redes sociais pra diferentes fins.

Essa análise da autora é positiva, pois permite iniciar discussões para além da representatividade e desigualdade existentes entre o Norte Global e o Sul Global, em relação ao perfil dos usuários das redes sociais (que são majoritariamente do Sul Global) mas também em relação a nacionalidade dos próprios conselheiros do FOB. Além disso, a alternativa de mais participação da sociedade civil por mais que não fosse a representação perfeita de toda sociedade, ainda contribui significativamente uma vez que as organizações da sociedade civil já tendem ser grupos organizados em prol de ações que visam beneficiar a sociedade em geral.

De maneira geral, o que senti falta no artigo foram de mais exemplos de aplicação em outros modelos de conselho ou instituições que incluam a participação da sociedade civil no processo de tomada de decisão e que sejam bem-sucedidos, como por exemplo conselhos de governança como o da Internet.

Citações

“The proposal, however, introduces a key divergence from the Habermasian or deliberative model in how it conceptualizes the purpose of participation. (...) Rather, the participatory process should illuminate the divergences among participants and whose perspectives are disfavored. Making this explicit also makes clear the need for those who were disfavored by the outcome of the process to retain access to contest that outcome in the future and maintain an open dialogue.” (P. 2)

“In the case of the FOB, this legitimization strategy works on two levels. The experts claim that they are compelled to rely on IHRL. And they try to show that IHRL can yield decisions when applied to a set of facts. Both moves present the experts as apolitical enforcers of the public’s will. The following sections illustrate how the Board uses IHRL to present itself as an objective decision-maker.” (P. 4)

“According to this view, the massive reach of speech on social media can transform harmful but tolerable expressions into intolerable instances of harm.” (P. 4)

“Put differently, even if companies were committed to reflecting users’ values, users around the world do not share the same preferences over what speech should be excluded from social media.” (P. 5)

“If that is true, Facebook might be more interested in appeasing the demands of certain civil society groups over others. Civil society located in more important markets or closer to more powerful regulators is indeed more influential than civil society in the global South.” (P. 8)

“Civil society is an interesting actor to involve more strongly in speech governance not because they perfectly represent the public. (P. 10)

Instead, among the available possibilities, the organized civil society is a site for hope because it is constituted by many actors advocating for what they believe benefits the society at large and representing multiple diverse and conflicting standpoints.” (P. 10)

Referências

JASANOFFJ, S. **The practices of objectivity in regulatory science**. In C. Camic, N. Gross, & M. Lamont (Eds.), *Social knowledge in the making* (pp. 307–338). Chicago: The University of Chicago Press, 2011.

KAYE, D. **Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression**. Geneva: United Nations, Human Rights Council. Report no. A/HRC/38/35. April 6. 2018.

FELDMAN, N. **A Supreme court for Facebook** (jan. 1, 2018). In *Global feedback and input on the Facebook oversight board for content decisions appendix*. <https://about.fb.com/wp-content/uploads/2019/06/oversight-Board-consultation-report-appendix.pdf>. 2018

DOUEK, e. **What kind of oversight board have you given us?** The University of Chicago Law Review Online. May 11 <https://lawreviewblog.uchicago.edu/2020/05/11/fb-oversight-Board-edouek/>. 2020

DOUEK, e. **The limits of international law in content moderation**. U.C. Irvine Journal of International, Transnational & Comparative Law, 6, 37–75. 2021

Kayyali, D., & York, J. **The Facebook oversight board is making good decisions—but does it matter?** Tech policy press. July 28 <https://techpolicy.press/the-facebook-oversight-Board-is-making-good-decisions-but-does-it-matter/>. 2021

Apêndice D – Ficha de Leitura: ARUN, Chinmayi. Facebook's

ARUN, Chinmayi. **Facebook's faces**. Case Western Reserve Journal of International Law, v. 53, n. 1, p. 1-36, 2021. Disponível em: <https://harvardlawreview.org/forum/no-volume/facebook-faces/>

Resumo

No artigo, "Facebook's Faces", a autora analisa a complexa rede de relações entre o Facebook e demais stakeholders como os Estados, a sociedade civil, seus próprios funcionários e o Oversight Board, destacando então que a empresa não é uma entidade monolítica. Para isso, é analisado como caso central a suspensão das contas de Donald Trump, expondo como o Conselho se encaixa no sistema de governança de conteúdo da plataforma. Arun argumenta que o Facebook possui "muitas faces", com equipes internas que, muitas vezes, têm objetivos desalinhados, impactando a moderação de conteúdo.

Assim, a autora apresenta o sistema de regras do Facebook que considera os padrões da comunidade, que se aplicam globalmente a todos os usuários, e as leis locais, que variam de país para país. A aplicação das leis locais, com pouca transparência além dos relatórios que a empresa divulga, não está sujeita à supervisão do Conselho. Além disso, ressalta que a aplicação desses padrões e leis é influenciada pelas pressões de estados e públicos, bem como pelas prioridades internas do Facebook, destacando que nem todos os estados ou públicos têm o mesmo poder de influência. Como exemplo, o artigo cita, que embora a Comissão Europeia tenha conseguido persuadir o Facebook a seguir um Código de Conduta voluntário sobre discurso de ódio, Mark Zuckerberg recusou convites para comparecer perante parlamentos de outros países, como Canadá e Índia.

Em conclusão, Arun sugere que o Conselho de Supervisão pode exercer influência ao se aliar a outros atores, explorando as tensões internas do Facebook e as diferentes prioridades de suas equipes. O artigo também enfatiza a necessidade de maior transparência e cooperação por parte da empresa para que o Conselho possa desempenhar seu papel de forma eficaz, contribuindo para uma melhor governança

do conteúdo online e para uma jurisprudência compartilhada globalmente e que visam garantir a equidade na moderação de conteúdo, dada a complexidade e a falta de transparência inerentes a esses sistemas.

Avaliação Crítica

Um aspecto que se destaca no artigo é a análise de como o Facebook se relaciona de maneira desigual com os diferentes atores. A autora demonstra com diferentes exemplos que nem todos os estados ou públicos têm o mesmo poder de influência sobre a empresa, que tende a priorizar seus interesses comerciais e de reputação ao lidar com questões de moderação de conteúdo. A citação dos casos da China e do Irã, onde o Facebook optou por não operar em vez de ceder às demandas dos governos locais, e a comparação com a postura da empresa em relação à União Europeia, que conseguiu persuadir o Facebook a seguir um "Código de Conduta" sobre discurso de ódio, ilustram essa disparidade de influência, assim por mais que seja explicado que o Facebook vai seguir suas próprias diretrizes e as leis locais, o quanto se estaria disposto a seguir leis locais é variável e com isso o alcance do Conselho, em decidir sobre alguns casos também se torna limitado.

Também é interessante como a autora descreve que as diferentes equipes e indivíduos dentro da empresa têm diferentes prioridades e níveis de influência, e como isso afeta a tomada de decisões sobre moderação de conteúdo, que é onde a maioria das decisões vão ser tomadas – e não com o Conselho. O sistema Cross Check, que concede um tratamento diferenciado a contas de alto perfil, é um exemplo que se podem aplicar regras de moderação de conteúdo de forma inconsistente e arbitrária. Por isso, eu concordo com a autora em relação a participação da sociedade civil nas decisões do Conselho de Supervisão ser ainda mais relevante, pois garante que as vozes dos usuários da plataforma sejam ouvidas e que suas solicitações sejam devidamente consideradas.

Apesar de mencionado, o artigo poderia explorar um pouco mais como o Conselho pode vir a ser apenas estratégico para garantir a legitimidade do Facebook junto aos Estados e ao mercado do que necessariamente um recurso como uma "Suprema Corte" para os usuários. Considerando que a capacidade de responsabilizar o

Facebook é limitada e que não pode atuar nos casos que sejam em relação a leis locais. A recusa do Facebook em cooperar plenamente com o Conselho no caso da suspensão de Donald Trump e o sistema de *Cross Check*.

Em última análise, "Facebook's Faces" é uma análise valiosa e oportuna da complexa paisagem da governança de conteúdo online. Ao destacar as tensões internas, as relações desiguais e as limitações do Conselho de Supervisão, Arun oferece uma perspectiva crítica que é essencial para entender os desafios e as oportunidades de regular as plataformas de mídia social de forma eficaz e justa.

Citações

"The Board was constructed as an independent entity with a limited sphere of influence and an apparent purpose to legitimize the company's self-regulation, which presumably included Facebook's decisionmaking about political leaders like Donald Trump. CEO Mark Zuckerberg described the Board as an entity that would make Facebook accountable to its community" (P. 238)

"The bifurcated system of rules described above looks simple at first glance — global Community Standards are between Facebook and its users, while local laws are about governments' relationships with their citizens, which Facebook respects." (P. 242)

"It is evident that Oversight Board cannot control the company. However, this is not to understate the Oversight Board's value or influence. Its expert advice has already improved Facebook's decisionmaking. But the answer to the more limited question of whether Facebook permits the Oversight Board to hold it accountable is no." (P. 262)

"It appears that the Oversight Board is settling into its role as a significant advisory body committed to engaging with Facebook, and showing the company and the rest of the world how Facebook can make better decisions about content. The Board has worked on underlining its independence and commitment to its own goals, such that I, for one, cannot think of it as one of Facebook's faces.." (P. 263)

Referências

GILLESPIE, T. **Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media**. New Haven: Yale University Press, 2018

GLOBAL NETWORK INITIATIVE. **GNI Assessment Toolkit**. 2021, p. 16, 67. Disponível em: <https://globalnetworkinitiative.org/wp-content/uploads/2021/11/AT2021.pdf>.

BLOCH-WEHBA, H. **Global Platform Governance and Private Power in the Shadow of the State**. *SMU Law Review*, v. 72, p. 27-71, 2019, p. 71.

DOUEK, E. **Facebook's "Oversight Board:" Move Fast with Stable Infrastructure and Humility**. *North Carolina Journal of Law & Technology*, v. 21, p. 1, 2019, p. 3.

KLONICK, K. **The New Governors: The People, Rules, and Processes Governing Online Speech**. *Harvard Law Review*, v. 131, p. 1598, 2018.

BALKIN, J. M. **Essay, Free Speech Is a Triangle**. *Columbia Law Review*, v. 8, p. 2011-2015-21, 2018.

KELLER, D. **Who Do You Sue? State and Platform Hybrid Power over Online Speech**. Hoover Institution Working Group on National Security, Technology, & Law, Aegis Series Paper No. 1902, 2019, p. 3-7. Disponível em: https://www.hoover.org/sites/default/files/research/docs/who-do-you-sue-state-and-platform-hybrid-power-over-online-speech_0.pdf.

LAND, M. K. **Against Privatized Censorship: Proposals for Responsible Delegation**. *Virginia Journal of International Law*, v. 60, p. 363, 2020, p. 380-88.

FRASER, N. **Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy**. *Social Text*, n. 25/26, 1990, p. 56, 57-61.

Apêndice F – Código de análise dos resultados

```
# === PRIMEIRO CÓDIGO (corrigido): importar libs, ler a aba "consolidado" e contar ===

import pandas as pd

import matplotlib.pyplot as plt

import unicodedata, re

from pathlib import Path

ARQUIVO = "20250816_PIBIC.xlsx"

ABA_ALVO = "consolidado" # <- corrigido

# Abrir o Excel e localizar a aba (case-insensitive)

xls = pd.ExcelFile(Path(ARQUIVO))

sheet = next((s for s in xls.sheet_names if s.strip().lower() == ABA_ALVO.lower()), None)

if sheet is None:

    raise ValueError(f"Aba '{ABA_ALVO}' não encontrada. Abas disponíveis: {xls.sheet_names}")

df = pd.read_excel(xls, sheet_name=sheet)

# Normalizar nomes de colunas para achar "Number of public comments" com segurança

def norm(s: str) -> str:

    s = unicodedata.normalize("NFKD", s)

    s = "".join(ch for ch in s if not unicodedata.combining(ch))

    s = s.strip().lower()

    s = re.sub(r"\s+", "", s)

    s = re.sub(r"[^a-z0-9]", "", s)

    return s

norm_map = {norm(c): c for c in df.columns}

key = "numberofpubliccomments"

if key not in norm_map:

    raise ValueError(f"Coluna 'Number of public comments' não encontrada. Colunas disponíveis")

col_comments = norm_map[key]

# Converter para numérico e contar

df["_n_comments"] = pd.to_numeric(df[col_comments], errors="coerce").fillna(0)
```

```

n_total = len(df)

n_com = int((df["_n_comments"] >= 1).sum())

n_sem = int((df["_n_comments"] < 1).sum())

p_com = (n_com / n_total * 100) if n_total else 0.0

p_sem = (n_sem / n_total * 100) if n_total else 0.0

print(f"Aba usada: {sheet}")

print(f"Total de decisões: {n_total}")

print(f"Com comentários públicos (>=1): {n_com} ({p_com:.1f}%)")

print(f"Sem comentários públicos (0 ou NaN): {n_sem} ({p_sem:.1f}%)")

# --- Gráfico rápido ---

labels = ["Com comentários", "Sem comentários"]

valores = [n_com, n_sem]

cores = ["#1f77b4", "#7f7f7f"] # azul, cinza

fig, ax = plt.subplots(figsize=(6,4))

bars = ax.bar(labels, valores, color=cores)

for i, (v, pct) in enumerate(zip(valores, [p_com, p_sem])):

    ax.text(i, v, f"{v} ({pct:.1f}%)", ha="center", va="bottom", fontsize=10, fontweight="bo

ax.set_title("Decisões que receberam comentários públicos")

ax.set_xlabel("")

ax.set_ylabel("Número de decisões")

plt.xticks(rotation=0)

plt.tight_layout() plt.show()

```

```

# --- Gráfico (sem borda, sem eixo Y, rótulos com mais espaço) ---

labels = ["Com comentários", "Sem comentários"]

valores = [n_com, n_sem]

pcts = [p_com, p_sem]

cores = ["#1f77b4", "#7f7f7f"]

fig, ax = plt.subplots(figsize=(6, 4))

bars = ax.bar(labels, valores, color=cores)

ymax = max(valores) if valores else 0

offset = max(1, 0.06 * ymax) # ↑ aumenta o espaço acima da barra

for i, (v, pct) in enumerate(zip(valores, pcts)):

    ax.text(i, v + offset, f"{v} ({pct:.1f}%",

ha="center", va="bottom", fontsize=10, fontweight="bold")

ax.set_title("Número de decisões que receberam comentários públicos")

ax.set_xlabel("") # sem texto no eixo X

ax.set_ylabel("") # sem texto no eixo Y

ax.get_yaxis().set_visible(False) # remove o eixo Y inteiro (linha + números)

# remove todas as bordas (spines)

for spine in ["top", "right", "left", "bottom"]:

    ax.spines[spine].set_visible(False)

ax.grid(False)

ax.set_ylim(0, ymax + offset * 3) # folga para os rótulos

plt.xticks(rotation=0)

plt.tight_layout()

plt.savefig("decisoões_comentarios_publicos.png", dpi=300, bbox_inches="tight")

plt.show()

```

```

# --- Gráfico: Número de casos que receberam comentários públicos por ano ---

import pandas as pd

import matplotlib.pyplot as plt

import unicodedata, re

# se já tiver 'df' carregado da aba "consolidado", use-o; caso contrário, descomente abaixo:

# df = pd.read_excel("20250816_PIBIC.xlsx", sheet_name="consolidado")

def norm(s: str) -> str:

    s = unicodedata.normalize("NFKD", s)

    s = "".join(ch for ch in s if not unicodedata.combining(ch))

    s = s.strip().lower()

    s = re.sub(r"^\s+", "", s)

    s = re.sub(r"[^a-z0-9_]", "", s)

    return s

colmap = {norm(c): c for c in df.columns}

col_comments_key = "numberofpubliccomments"

col_date_key = "decision_date_ajustado_ruth"

if col_comments_key not in colmap:

    raise ValueError(f"Coluna 'Number of public comments' não encontrada. Colunas: {list(df.

if col_date_key not in colmap:

    raise ValueError(f"Coluna 'decision_date_ajustado_ruth' não encontrada. Colunas: {list(d

# filtrar apenas casos com comentários (>=1) e com data válida

df2 = df.copy()

df2["_n_comments"] = pd.to_numeric(df2[colmap[col_comments_key]], errors="coerce").fillna(0)

df2["_dt"] = pd.to_datetime(df2[colmap[col_date_key]], errors="coerce", dayfirst=True)

df2 = df2[(df2["_n_comments"] >= 1) & df2["_dt"].notna()].copy()

# agrupar por ano

df_year = (df2.assign(year=df2["_dt"].dt.year)

.groupby("year")

.size()

.reset_index(name="n_casos")

```

```
.sort_values("year"))

# plot

plt.close('all')

fig, ax = plt.subplots(figsize=(8, 5))

bars = ax.bar(df_year["year"].astype(int).astype(str), df_year["n_casos"], color="#1f77b4")

# rótulos no topo (número de casos)

ymax = df_year["n_casos"].max() if len(df_year) else 0

offset = max(1, 0.06 * ymax)

for rect, v in zip(bars, df_year["n_casos"]):

    ax.text(rect.get_x() + rect.get_width()/2, v + offset, f"{int(v)}",

            ha="center", va="bottom", fontsize=10, fontweight="bold")

ax.set_title("Número de casos que receberam comentários públicos por ano")

ax.set_xlabel("Ano")

ax.set_ylabel("") # sem texto no eixo Y

ax.get_yaxis().set_visible(False) # remove eixo Y

# remover bordas (spines)

for spine in ["top", "right", "left", "bottom"]:

    ax.spines[spine].set_visible(False)

ax.set_ylim(0, ymax + offset * 3)

plt.xticks(rotation=0)

plt.tight_layout()

# salvar antes de mostrar (opcional)

plt.savefig("casos_comentarios_publicos_por_ano.png", dpi=300, bbox_inches="tight")

plt.show()
```

```

bins = [1,2,3,5,10,20,50,100, np.inf]

labels = ["1", "2", "3-5", "6-10", "11-20", "21-50", "51-100", "100+"]

freq = pd.cut(s, bins=bins, labels=labels, include_lowest=True, right=True).value_counts()

freq = freq.reindex(labels) # mantém ordem das faixas

perc = (freq / freq.sum()) * 100).round(1)

plt.close('all')

fig, ax = plt.subplots(figsize=(9, 5))

bars = ax.bar(freq.index.astype(str), freq.values, color="#1f77b4")

ymax = freq.values.max() if len(freq) else 0

offset = max(1, 0.06 * ymax)

for i, (v, p) in enumerate(zip(freq.values, perc.values)):
    ax.text(i, v + offset, f"{int(v)} ({p}%)",
           ha="center", va="bottom", fontweight="bold")

ax.set_title("Distribuição de casos por faixa de nº de comentários (≥1)")

ax.set_xlabel("Faixa de comentários por caso")

ax.set_ylabel(""); ax.get_yaxis().set_visible(False)

for sp in ["top", "right", "left", "bottom"]:
    ax.spines[sp].set_visible(False)

ax.set_ylim(0, ymax + offset*3)

plt.xticks(rotation=0)

plt.tight_layout()

plt.savefig("dist_comentarios_por_caso_faixas.png", dpi=300, bbox_inches="tight")

plt.show()

```

```

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

from pathlib import Path

from scipy.stats import chi2_contingency, fisher_exact

from sklearn.metrics import cohen_kappa_score

ARQ = "20250817_PIBIC.xlsx"

SHEET_CONS = "consolidado"

SHEET_COMM = "Comentarios"

MIN_COMMENTS = 1 # altere para 5 ou 3 se quiser um filtro de robustez

# --- helpers ---

def clean_label(x):

    """Normaliza rótulos textuais para 'keep' ou 'remove' (remove ponto final, ignora NaN)."""

    if pd.isna(x):

        return np.nan

    t = str(x).strip().lower().rstrip(".")

    mapping = {

        "keep the content": "keep",

        "remove the content": "remove",

        # sinônimos comuns (se aparecerem):

        "restore the content": "keep",

        "uphold the removal": "remove",

        "keep": "keep",

        "remove": "remove",

    }

    if t in {"", "nan"}:

        return np.nan

    return mapping.get(t, t)

def find_id_col(df):

    """Acha a coluna de ID tolerando variações."""

```

```

names = {c.strip().lower(): c for c in df.columns}

for k in ("id", "codigo", "código", "code", "case_id"):

if k in names:

return names[k]

raise ValueError(f"Não achei coluna de ID. Colunas: {list(df.columns)}")

# --- carregar dados ---

xls = pd.ExcelFile(Path(ARQ))

cons = pd.read_excel(xls, sheet_name=SHEET_CONS)

def find_id_col(df):

"""Acha a coluna de ID tolerando variações."""

names = {c.strip().lower(): c for c in df.columns}

for k in ("id", "codigo", "código", "code", "case_id"):

if k in names:

return names[k]

raise ValueError(f"Não achei coluna de ID. Colunas: {list(df.columns)}")

# --- carregar dados ---

xls = pd.ExcelFile(Path(ARQ))

cons = pd.read_excel(xls, sheet_name=SHEET_CONS)

comm = pd.read_excel(xls, sheet_name=SHEET_COMM)

# --- detectar colunas de ID ---

cons_id_col = find_id_col(cons)

comm_id_col = find_id_col(comm)

print(f"[OK] IDs -> consolidado: '{cons_id_col}' | Comentarios: '{comm_id_col}'")

# --- padronizar IDs e rótulos ---

cons["id"] = cons[cons_id_col].astype(str).str.strip()

comm["id"] = comm[comm_id_col].astype(str).str.strip()

cons["_ob"] = cons["What was OB decision?"].apply(clean_label)

comm["_cls"] = comm["classification_ajustado"].apply(clean_label)

valid = ["keep", "remove"]

```

```

cons_v = cons[cons["_ob"].isin(valid)].copy()

comm_v = comm[comm["_cls"].isin(valid)].copy()

# --- agregar comentários por decisão ---

counts = (

    comm_v.groupby(["id", "_cls"])

    .size()

    .unstack(fill_value=0)

    .reindex(columns=valid, fill_value=0)

    .astype(int)

)

def maj_margin(row):

    k, r = row["keep"], row["remove"]

    tot = k + r

    if tot < MIN_COMMENTS:

        return pd.Series({"majority": np.nan, "margin": np.nan, "n_labeled": int(tot)})

    if k == r:

        return pd.Series({"majority": np.nan, "margin": 0.0, "n_labeled": int(tot)})

    # empate

    winner = "keep" if k > r else "remove"

    margin = abs(k - r) / tot

    return pd.Series({"majority": winner, "margin": margin, "n_labeled": int(tot)})

maj = counts.apply(maj_margin, axis=1)

by_case = counts.join(maj).reset_index()

# --- juntar com OB e filtrar válidos ---

res = by_case.merge(cons_v[["id", "_ob"]], on="id", how="inner")

anal = res[res["majority"].isin(valid) & res["_ob"].isin(valid)].copy()

anal["converge"] = (anal["majority"] == anal["_ob"])

# --- métricas principais ---

n_cases = len(anal)

n_conv = int(anal["converge"].sum())

```

```

p_conv = (n_conv / n_cases * 100) if n_cases else np.nan

print(f"\nCasos com maioria de comentários (n≥{MIN_COMMENTS}) e OB válido: {n_cases}")

print(f"Convergência (OB = maioria): {n_conv} casos ({p_conv:.1f}%)")

tab = (
    pd.crosstab(anal["majority"], anal["_ob"])
    .reindex(index=valid, columns=valid, fill_value=0)
)

print("\nContingência (maioria x OB):\n", tab)

if tab.values.sum() > 0:
    chi2, p, dof, exp = chi2_contingency(tab.values)

    # 2x2

    phi = np.sqrt(chi2 / tab.values.sum())

    OR, p_fisher = fisher_exact(tab.values) # robusto p/ células pequenas

    kappa = cohen_kappa_score(anal["majority"], anal["_ob"])

    print(f"\nQui-quadrado: {chi2:.3f} | p-valor: {p:.5f} | gl: {dof} | Phi: {phi:.3f}")

    print(f"Fisher exact: OR={OR:.3f} | p={p_fisher:.5f}")

    print(f"Cohen's kappa: {kappa:.3f}")

# --- gráfico (opcional) ---

vals = anal["converge"].value_counts().reindex([True, False]).fillna(0).astype(int)

labels = ["Convergência", "Divergência"]

cores = ["#1f77b4", "#7f7f7f"]

plt.close('all')

fig, ax = plt.subplots(figsize=(6, 4))

bars = ax.bar(labels, vals.values, color=cores)

ymax = vals.values.max() if len(vals) else 0

offset = max(1, 0.06*ymax)

tot = vals.sum()

for i, v in enumerate(vals.values):

    pct = v/tot*100 if tot else 0

    ax.text(i, v + offset, f"{int(v)} ({pct:.1f}%)", ha="center", va="bottom", fontweight="b

```

```

ax.set_title("OB seguiu a maioria dos comentários?")

ax.set_xlabel(""); ax.set_ylabel(""); ax.get_yaxis().set_visible(False)

for s in ["top", "right", "left", "bottom"]:

ax.spines[s].set_visible(False)

ax.set_ylim(0, ymax + offset*3)

plt.tight_layout()

plt.show()

# (opcional) exportar detalhe por caso

cols_out = ["id", "_ob", "majority", "margin", "n_labeled", "keep", "remove"]

anal[cols_out].sort_values("margin", ascending=False).to_excel("convergencia_OB_por_caso.xls

print("\nArquivo salvo: convergencia_OB_por_caso.xlsx")

```

```

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

from scipy.stats import pearsonr

from sklearn.linear_model import LogisticRegression

# --- 1) Construir as variáveis numéricas ---

# X = proporção de comentários pró-keep (entre 0 e 1)

X = (anal["keep"] / (anal["keep"] + anal["remove"])).astype(float)

# Y = decisão do OB (keep=1, remove=0)

Y = (anal["_ob"] == "keep").astype(int)

# correlação ponto-biserial (equivalente ao Pearson com Y binário)

r, p = pearsonr(X, Y)

# --- 2) Scatter com jitter + curva logística ---

rng = np.random.default_rng(42)

Y_jit = Y + rng.uniform(-0.04, 0.04, size=len(Y)) # jitter pequeno só p/ visual

plt.close('all')

fig, ax = plt.subplots(figsize=(7.5, 5))

# pontos

ax.scatter(X, Y_jit, s=40, alpha=0.85, color="#1f77b4")

```

```

# linha vertical no limiar 0.5 (metade dos comentários pró-keep)

ax.axvline(0.5, linestyle="--", linewidth=1)

# curva logística (tendência)

clf = LogisticRegression(solver="lbfgs")

clf.fit(X.values.reshape(-1, 1), Y.values)

grid = np.linspace(0, 1, 300).reshape(-1, 1)

proba = clf.predict_proba(grid)[:, 1]

ax.plot(grid, proba, linewidth=2)

# eixos e rótulos

ax.set_title("Decisão do OB vs. proporção de comentários para manter")

ax.set_xlabel("Proporção de comentários para manter (0–1)")

ax.set_ylabel("")

ax.set_yticks([0, 1])

ax.set_yticklabels(["OB Remover", "OB Manter"])

# estilo (sem borda sup./dir. e sem eixo Y numérico)

for s in ["top", "right"]:

    ax.spines[s].set_visible(False)

ax.grid(False)

ax.set_xlim(-0.02, 1.02)

ax.set_ylim(-0.2, 1.2)

# anotação da correlação

ax.text(0.02, 1.12, f"r = {r:.2f} (p = {p:.3f})", fontsize=10)

plt.tight_layout()

# plt.savefig("scatter_correlacao_OB_vs_prop_keep.png", dpi=300, bbox_inches="tight")

plt.show()

```

```

from scipy.stats import pearsonr, norm

import numpy as np

r, p = pearsonr(X, Y)

r2 = r**2

# IC95% de r (transformação z de Fisher)

```

```

n = len(X)

z = np.arctanh(r)

se = 1/np.sqrt(n-3)

zcrit = norm.ppf(0.975)

lo, hi = np.tanh(z - zcrit*se), np.tanh(z + zcrit*se)

print(f"r = {r:.2f}, r² = {r2:.2f}, p = {p:.3f}, IC95% = [{lo:.2f}, {hi:.2f}]")

```

```

import seaborn as sns

import matplotlib.pyplot as plt

# --- mapa de tradução (adicione mais se aparecerem outros termos) ---

trad = {

"Freedom of expression": "Liberdade de expressão",

"Elections": "Eleições",

"Governments": "Governo",

"Children": "Crianças",

"Politics": "Política",

"Facebook": "Facebook",

"Art": "Arte",

"Discrimination": "Discriminação",

"Health": "Saúde",

"Journalism": "Jornalismo",

"Outros": "Outros",

}

# pct: dataframe % por ano (linhas = anos, colunas = temas)

pct_pt = pct.rename(columns=trad)

plt.close('all')

fig, ax = plt.subplots(figsize=(12, 6))

sns.heatmap(

pct_pt.T, cmap="Blues", cbar_kws={"label": "% no ano"},

annot=True, fmt=".0f", ax=ax

)

```

```

ax.set_title("Participação por tema principal — % dentro de cada ano")

ax.set_xlabel("Ano")

ax.set_ylabel("Tema")

# "Legenda"/nota explicativa abaixo do gráfico

plt.figtext(

0.5, -0.03,

"Nota: visualização mantém as 10 categorias mais frequentes; as demais foram agregadas e

ha="center", fontsize=9

)

plt.tight_layout()

plt.savefig("tema_principal_heatmap_pt.png", dpi=300, bbox_inches="tight")

plt.show()

```

```

import pandas as pd

import matplotlib.pyplot as plt

import numpy as np

# ===== Parâmetros =====

ARQ = "20250817_PIBIC.xlsx"

ABA = "consolidado"

COL_ID = "id"

COL_COUNTRY = "country"

DROP_INDEFINIDO = True # trocar para False se quiser manter a barra "Indefinido"

# ===== 1) Carregar =====

xls = pd.ExcelFile(ARQ)

sheet = next((s for s in xls.sheet_names if s.strip().lower() == ABA.lower()), None)

if sheet is None:

raise ValueError(f"Aba '{ABA}' não encontrada. Abas: {xls.sheet_names}")

df = pd.read_excel(xls, sheet_name=sheet)

colmap = {c.strip().lower(): c for c in df.columns}

for req in [COL_ID, COL_COUNTRY]:

if req.lower() not in colmap:

```

```

raise ValueError(f"Coluna '{req}' não encontrada. Colunas: {df.columns.tolist()}")

c_id = colmap[COL_ID.lower()]

c_cty = colmap[COL_COUNTRY.lower()]

df = df[[c_id, c_cty]].copy()

df[c_cty] = df[c_cty].fillna("").astype(str)

# ===== 2) Explodir países por decisão =====

df["__tokens"] = df[c_cty].apply(lambda s: [t.strip() for t in s.split(",") if t.strip()])

expl = df.explode("__tokens").rename(columns={"__tokens": "country_raw"})

expl["country_raw"] = expl["country_raw"].astype(str).str.strip()

expl = expl.drop_duplicates(subset=[c_id, "country_raw"])

# ===== 3) Mapear país -> continente =====

NA = {
    "united states", "usa", "u.s.", "us", "u.s.a.", "canada", "mexico",
    "guatemala", "belize", "honduras", "el salvador", "nicaragua", "costa rica", "panama",
    "cuba", "haiti", "dominican republic", "puerto rico", "jamaica", "trinidad and tobago",
    "bahamas", "barbados", "grenada", "saint lucia", "saint kitts and nevis", "antigua and barbud
}

SA = {
    "brazil", "argentina", "uruguay", "paraguay", "chile", "bolivia", "peru", "ecuador",
    "colombia", "venezuela", "guyana", "suriname",
}

EU = {
    "united kingdom", "england", "scotland", "wales", "northern ireland", "ireland", "france",
}

EU = {
    "spain", "portugal", "italy", "netherlands", "belgium",
    "switzerland", "austria", "poland", "czech republic", "slovakia",
    "hungary", "romania", "bulgaria", "greece", "sweden",
    "norway", "denmark", "iceland", "ukraine", "russia",
    "latvia", "lithuania", "estonia", "croatia", "serbia",
    "bosnia and herzegovina", "slovenia", "north macedonia",
}

```

"albania", "moldova"

}

OC = {

"australia", "new zealand", "papua new guinea", "fiji",

"samoa", "tonga", "vanuatu"

}

AF = {

"south africa", "egypt", "nigeria", "ethiopia", "sudan",

"south sudan", "somalia", "kenya", "tanzania", "uganda",

"rwanda", "burundi", "ghana", "senegal", "libya", "eritrea",

"côte d'ivoire", "morocco", "algeria", "tunisia", "niger",

"guinea-bissau"

}

AS = {

"india", "china", "japan", "south korea", "korea, republic of",

"thailand", "myanmar", "cambodia", "laos", "vietnam",

"philippines", "indonesia", "malaysia", "singapore",

"sri lanka", "nepal", "bangladesh", "pakistan", "afghanistan",

"mongolia"

}

ME = {

"israel", "palestinian territories", "palestinian territory",

"iran", "iraq", "syria", "lebanon", "jordan", "saudi arabia",

"yemen", "oman", "qatar", "bahrain", "kuwait",

"united arab emirates"

}

```
def norm_country_to_continent(token: str) -> str:
```

```
    t = token.strip()
```

```
    if not t:
```

```
        return "Indefinido"
```

```
    low = t.lower()
```

```
    syn = {
```

```
        "u.k.": "united kingdom",
```

```
        "uk": "united kingdom",
```

```
        "u.s.": "united states",
```

```
        "u.s.a.": "united states",
```

```
        "usa": "united states",
```

```
        "russian federation": "russia",
```

```
        "ivory coast": "côte d'ivoire",
```

```
        "czech republic": "czechia",
```

```
        "myanmar (burma)": "myanmar",
```

```
        "burma": "myanmar",
```

```
        "korea, republic of": "south korea",
```

```
    }
```

```
    low = syn.get(low, low)
```

```
    if low in NA:
```

```
        return "América do Norte"
```

```
    if low in SA:
```

```
        return "América do Sul"
```

```
    if low in EU:
```

```
        return "Europa"
```

```
    if low in OC:
```

```

    return "Oceania"

if low in AF:

    return "África"

if low in ME:

    return "Oriente Médio"

if low in AS:

    return "Ásia"

return "Indefinido"

expl["_continente"] = expl["country_raw"].map(norm_country_to_continent)

# ===== 4) Contar decisões por continente (sem filtro) =====

# Conta cada decisão no(s) continente(s) correspondente(s), no máximo 1 vez por continente.

by_cont = (

    expl.dropna(subset=["_continente"])

    .drop_duplicates(subset=[c_id, "_continente"])

    .groupby("_continente")[c_id]

    .nunique()

    .sort_values(ascending=False)

    .rename("n_casos")

    .reset_index()

)

if DROP_INDEFINIDO:

    by_cont = by_cont[by_cont["_continente"] != "Indefinido"]

print("Casos por continente (sem filtro):")

print(by_cont)

```