

Public Comment Appendix for

2023-001-FB-UA

Case number

Case description

On January 3, 2023, two days after Luiz Inácio Lula da Silva had been sworn in as **Brazil's president**, a Facebook user posted a video with a caption in Portuguese. The caption includes a call to "besiege" Brazil's congress as "the last alternative." The video shows part of a speech given by a prominent Brazilian general and supporter of Lula's electoral opponent, in which he calls for people to "hit the streets" and "go to the National Congress... [and the] Supreme Court." A sequence of images follows the general's speech, including one of a fire raging in the Three Powers Plaza in Brasília, which houses Brazil's presidential offices, Congress, and Supreme Court. Text overlaying the image reads, "Come to Brasília! Let's Storm it! Let's besiege the three powers." Text overlaying another image reads "we demand the source code," a slogan that protestors have used to question the reliability of Brazil's electronic voting machines. The video was played over 18,000 times, was not shared, and was reported seven times.

Mr. Lula da Silva's swearing-in had been accompanied by civil unrest, including protests and roadblocks. On January 8, more than a thousand supporters of former president Jair Bolsonaro broke into the National Congress, Supreme Court, and presidential offices, intimidating the police and destroying property. Meta designated Brazil a temporary high-risk location ahead of the country's October 2022 general election, and has been removing content "calling for people to take up arms or forcibly invade ...federal buildings" as a consequence. Meta only announced it had done so on January 9.

On the same day the content was posted, a user reported it for violating Meta's Violence and Incitement Community Standard, which prohibits calls to "forcibly enter locations ... where there are temporary signals of a heightened risk of violence or offline harm." In total, four users reported the content seven times between January 3 and January 4. Following the first report, the content was reviewed by a human reviewer and found **not to violate Meta's policies**. The user appealed the decision, but it was upheld by a second human reviewer. The next day, the other six reports were reviewed by five different moderators, all of whom found **that it did not violate Meta's policies**. The content was not escalated to policy or subject matter experts for additional review.

One of the users who had reported the content appealed Meta’s decision to the Oversight Board. In their appeal to the Board, they link the content’s potential to incite violence to the movement of people in Brazil “who do not accept the results of elections.”

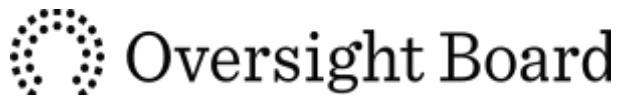
The Board selected this case to examine how Meta moderates election-related content, and how it is applying its Crisis Policy Protocol in a designated “temporary high-risk location.” Meta developed the Protocol in response to the Board’s recommendation in the “Former President Trump’s suspension” case. This case falls within the Board’s “Elections and civic space” priority.

As a result of the Board selecting this case, Meta determined that its repeated decisions to leave the content on Facebook were in error. Because at-scale reviewers do not record their reasons for making decisions, the company does not have further information about why they found the content did not violate its policies in this case. On January 20, 2023, Meta removed the content, issued a strike **against the content creator’s account, and applied a feature-limit**, preventing them from creating new content.

The Board would appreciate public comments that address:

- **The political situation in Brazil in advance of October’s election, and how it shifted between October 2022 and January 8, 2023.**
- The relationship between political violence, election denialism, and calls for offline mobilization on social media.
- **When Meta’s election integrity efforts should begin and end, and what criteria should guide decisions about those timeframes, particularly as they relate to transitions of power.**
- How Meta should distinguish between legitimate political organizing and harmful coordinated action.
- How Meta should treat content attacking or delegitimizing democratic institutions and processes.

In its decisions, the Board can issue policy recommendations to Meta. While recommendations are not binding, Meta must respond within 60 days. As such, the Board welcomes public comments proposing recommendations that are relevant to this case.



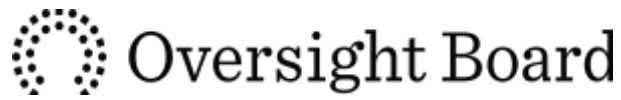
Public Comment Appendix for
2023-001-FB-UA
Case number

The Oversight Board is committed to bringing diverse perspectives from third parties into the case review process. To that end, the Oversight Board has established a public comment process.

Public comments respond to case descriptions based on the information provided to the Board by users and Facebook as part of the appeals process. These case descriptions are posted before panels begin deliberation to provide time for public comment. As such, case descriptions reflect neither the Board's assessment of a case, nor the full array of policy issues that a panel might consider to be implicated by each case.

To protect the privacy and security of commenters, comments are only viewed by the Oversight Board and as detailed in the [Operational Privacy Notice](#). All commenters included in this appendix gave consent to the Oversight Board to publish their comments. For commenters who did not consent to attribute their comments publicly, names have been redacted. To withdraw your comment, please email contact@osbadmin.com.

To reflect the wide range of views on cases, the Oversight Board has included all comments received except those clearly irrelevant, abusive or disrespectful of the human and fundamental rights of any person or group of persons and therefore violating the [Terms for Public Comment](#). Inclusion of a comment in this appendix is not an endorsement by the Oversight Board of the views expressed in the comment. The Oversight Board is committed to transparency and this appendix is meant to accurately reflect the input we received.



Public Comment Appendix for

2023-001-FB-UA

Case number

19

Number of Comments

Regional Breakdown

1	1	0	12
Asia Pacific & Oceania	Central & South Asia	Europe	Latin America & Caribbean
2	0	3	
Middle East and North Africa	Sub-Saharan Africa	United States & Canada	

2023-001-FB-UA

PC-10984

Latin America

Case number

Public comment number

Region

Rafael

Rubio

Commenter's first name

Commenter's last name

Commenter's preferred language

Observatorio de la desinformación.
Universidad Complutense

Organization

Yes

Response on behalf of organization

Short summary provided by the commenter

The observation report on disinformation in the electoral campaign, carried out as an electoral observation mission at the request of the TSE, can serve as a legal framework and social and political context for the case subject to review.

Full Comment

The observation report on disinformation in the electoral campaign, carried out as an electoral observation mission at the request of the TSE, can serve as a legal framework and social and political context for the case subject to review.

Link to Attachment

[PC-10984.pdf](#)

2023-001-FB-UA

PC-11000

Latin America

Case number

Public comment number

Region

FLORA

ARDUINI

English

Commenter's first name

Commenter's last name

Commenter's preferred language

Ekō

Yes

Organization

Response on behalf of organization

Short summary provided by the commenter

Our comment is to directly challenge Meta's clear inaction ahead of Brazil's 2022 election; lack of investment in both human and AI moderation; the applicability of its own T&Cs; lack of transparency; direct contribution to the distribution of content containing hate speech, disinformation, violence and military coup calls; and monetization over illegal content via Ads and algorithmic recommendation of such content.

Full Comment

Ekō published four investigative reports focused on Brazilian elections, and on several occasions, we identified how Meta's own tools and algorithms were directly aiding those seeking to build support for and orchestrate a military coup, including by directing users to content raising doubt about the integrity of the election. The reports provide concrete and damning evidence that Meta's platforms continue to fuel Brazil's own 'Stop the Steal' movement.

Our comment is to directly challenge Meta's clear inaction ahead of Brazil's 2022 election; lack of investment in both human and AI moderation; the applicability of its own T&Cs; lack of transparency; direct contribution to the distribution of content containing hate speech, disinformation, violence and military coup calls; and monetization over illegal content via Ads and algorithmic recommendation of such content.

Link to Attachment

[PC-11000.pdf](#)

2023-001-FB-UA

PC-11007

Asia Pacific and Oceania

Case number

Public comment number

Region

Dr Harry

Melkonian

English

Commenter's first name

Commenter's last name

Commenter's preferred language

Macquarie University Law School -
Media Law Students

No

Organization

Response on behalf of organization

Short summary provided by the commenter

Political speech attracts the highest level of protection because it is essential for effective democratic government. Political speech may not always be popular and the causes espoused vary from the commendable to the loathsome. The protections accorded political speech should not depend on the evaluation of the merits of the cause. However, speech or expression transcends the bounds of legitimacy when dangerous violence to persons or property is imminent. In the case of the situation in Brasilia, given the context of buildings alight, restraint or moderation of speech may be justified. Inasmuch as the speech in question clearly had political overtones, the Oversight Board must act with transparency and furnish a detailed explanation.

Full Comment

This submission addresses two of the issues raised by the Board:

1. How Meta should distinguish between legitimate political organizing and harmful coordinated action.
2. How Meta should treat content attacking or delegitimizing democratic government.

Political speech attracts the highest level of protection because it is essential for effective democratic government. Political speech may not always be popular and the causes espoused vary from the commendable to the loathsome. The protections accorded political speech should not depend on an evaluation of the merits of the cause. However, speech or expression transcends the bounds of legitimacy when dangerous violence to persons or property is imminent. In the case of the situation in Brasilia, given the context of buildings alight, restraint or moderation of speech

may be justified. Inasmuch as the speech in question clearly had political overtones, the Oversight Board must promote transparency and explain why this particular speech would be curtailed.

This submission is being made through the collective efforts of the over 200 students currently enrolled in Media Law at Macquarie University Law School, Sydney Australia. This submission consists of a synthesis of the drafts prepared by students working collaboratively. It should be noted that these students have already completed a comparative study of freedom of expression theory and laws.

1. How Meta should distinguish between legitimate political organizing and harmful coordinated action.

Limitations on freedom of political expression should be the exception and, the exception should be narrow. Speech, used in its broadest sense, to include all expressive acts, should be permitted even if unpopular. Issues-related speech should nearly always be allowed whereas there should be much stricter scrutiny about speech directed against individuals or groups, especially vulnerable minorities.

Provocative speech should be assessed based on its ability to incite action that will result in actual harm and tangible physical violence to persons or property.

Sometimes speech appears to support demonstrably provable falsehoods – such as abolition of naval forces because the earth is flat or no more trips to the moon because it is made of green cheese. Nevertheless, this speech should not be suppressed because it is political and highly unlikely to result in actual physical harm. However, even such eccentric speech can cross the line when the flat-earthlers urge violence against sailors or sabotage of naval installations. The challenge is to delineate where exactly that line falls.

As media law students conversant with freedom of expression theory, we conclude that bright lines in the sand are elusive and should not be attempted. Rather, the Oversight Board must consider not only the content but also the size of the audience and the comments the Facebook entry has encouraged. Of course, the actual situation on the ground may be the most important factor. That is, if there are buildings already alight, further calls for marches and sieges may be legitimately suppressed. On the other hand, a post calling for a march on the capital to start the revolution can be quite permissible if it occurs in a calm political environment. If the speech incites hate or violence against an individual or group, then Meta should not provide a forum. Furthermore, even if the speech is more clearly issue-related and clearly political, then the environmental factors become the dominant factors. Under this standard, the flat-earth society may demand abolition of the navy but, as buildings are aflame in Brasilia, calls for mass protest in that city should be curtailed, at least until the situation cools.

2. How Meta should distinguish between legitimate political organizing and harmful coordinated action.

Having started with the proposition that suppression of political speech should be a narrowly construed exception, it is then essential that Meta provides very clear parameters and is fully transparent when this censorship authority is exercised. Although the term is overused in modern discourse, transparency is key. If political speech is being suppressed, all parties must fully understand the specific circumstances that justified such extreme action.

Political speech that directs hate or violence against an individual or group, especially vulnerable groups, can be suppressed without further explanation. Speech that does not fall within that category should generally be permitted unless the environment makes the expression likely to cause imminent violence. That is, it is not so much the literal words but how those words are likely to be received that is the boundary between the permissible and the prohibited.

As an example, residents of Norway carrying signs that say ‘Death to Norway’s Pharaoh’, while calling for violence, are so without meaning as not to constitute a threat to persons or property, and should be allowed. However, those same words might pose different connotations and risks if espoused in Egypt. But, maybe not as Cleopatra was the last Pharaoh. But, maybe yes, if the current President of Egypt was being called a would-be Pharaoh.

What this points out is the essential role played by transparency. It’s not that different rules apply to Egypt as opposed to Norway but that different environmental situations exist that might justify censorship in the former. From Meta’s standpoint, the rule should be full and clear explanation of how the speech, regardless of literal content, might cause harm within its intended audience.

In the context of what occurred in Washington DC a year earlier and, within the incendiary climate of Brasilia – as some government buildings had already been torched – Meta needed expand upon about the fragility of civil order and the urgency for restraint. It was not the literal words spoken but the environment that surrounded those words that justify removal.

[Link to Attachment](#)

No Attachment

2023-001-FB-UA

PC-11008

Latin America

Case number

Public comment number

Region

Thales

Bueno

English

Commenter's first name

Commenter's last name

Commenter's preferred language

London School of Economics and
Political Science - LSE

No

Organization

Response on behalf of organization

Short summary provided by the commenter

This public comment is divided in two parts and draws upon the four bullet points asked in the case. Part 1 provides a contextual analysis of Brazilian elections. How has it changed after Lula's victory? What were key mistakes committed by Meta during the period? In Part 2 I write suggestions on how Meta should review its crisis protocol, considering 3 main fields: 1) Human Rights and international Law as the bedrock of content policies 2) pre-elections risk approach 3) During elections risk approach.

Part 2 will be sent in another comment.

Full Comment

Introductory remarks:

- 1) Meta did the right choice removing the content and applying penalties for the user;
- 2) I am an **ex-vendor from Meta's Governments, Politicians and Non-Profits (GPN)**, working directly in the elections plan of 2022 and this comment is not a personal revenge, but a reflection exercise.

Part 1: Brazilian context and the consequences of online actions threatening democracies.

As notorious and published over national and international media outlets, Brazilian elections were highly polarized between the former president Lula and the current one Jair Bolsonaro. The narrative of Jair Bolsonaro and its allies questioning the legitimacy of elections were in the center of the discussions around elections integrity.

The narrative of fraud, the call against the supreme court and the fear of political violence justifies a correct decision of Meta to consider Brazilian elections as “high risk”;

The result of the second run was tight. Lula beat Bolsonaro by a difference of 2 million votes, the tightest elections in the history of Brazilian democracy. Adding this tight difference to the narrative of fraud, as noticed by Agencia Publica on 02/11, Bolsonaro’s supporters started a strategy of resistance and denial, from online to offline. Even with the international recognition from different nations and the start of the government transition in a peaceful manner, online and offline campaigns questioning the results were active, gaining traction.

As disclosed by the of Agencia Publica, at the online level international actors such as Steve Bannon and Matthew Tyrmand tuned a narrative of fraud, organizing it by the hashtags #BrazillianSpring and #Braziawasstolen. At the offline level, Bolsonaro’s supporters blocked highways over 23 states and started a camping wave in front of the regional army center, asking for military intervention under the narrative of fraud. Even with the apparent pacific environment over these pro-militar campings, the narrative presented on social networks was a clear conspiracy theory and a call for military intervention. According to a survey made by CNN, there were 101 campings around the country before the episode of January 8th. Lula was sworn in a peaceful ceremony on January 1st and it looked like the high-risk elections in Brazil had ended. However, as published in the report from the initiative Digital Democracy, from January 1st the narrative under closed groups over WhatsApp, Facebook and mostly Telegram started to mobilize a riot, or the Selma’s party in the capital Brasilia that would happen on January 8th.

Following the report available on this article from Núcleo, the strategy disclosed over these groups and raised on Telegram were: 1) Invasion of the public buildings from Executive, Legislative and Supreme Court 2) Blocking oil refineries 3) Keep the campings alive.

The results of this “Capitolio Tropical” were intensively published and this public comment will not touch on the actions taken by the state.

Key mistakes and suggestions on what should have been done differently:

1. Content policies & Moderation: Concerning the question on how the context shifted after the end of the election, the key point was that even as a minority, from offline to online, people were acting against the integrity of the electoral process. Brazilian elections were recognized as fair at national and international level. Thus, users questioning the legitimacy of the result acted intentionally to spread misinformation. By doing so, Meta content moderation team could have applied community standards of dangerous individuals and organizations, but it seems that there was no clear guidance over this matter and content moderators have limited discretion over this kind of cases. As an example, as Youtube did print attached), Meta should have updated and published its new community guidelines in clear words right after the results were announced by the electoral supreme court.

2. Risk analysis: As reported by Netlab in the same article from Núcleo, Telegram was central to the mobilization of the "Capitolio Tropical". Therefore, a cross platform analysis would be necessary to measure the overall pulse of the political debate and determine the actions of the company. In other words, political speeches cross different platforms with different shapes and Meta has more than 20 million users over 3 products in Brazil. A broader risk analysis would be more accurate to understand the pulse of the political environment across the major platforms in Brazil.

Links:

<https://apublica.org/2023/01/festa-da-selma-bolsonaristas-usan-termino-de-uso-militar-para-coordinar-invasion-en-bsb/>

<https://nucleo.jor.br/curtas/2023-02-08-telegram/>

<https://support.google.com/youtube/answer/10835034?hl=pt-BR#zippy=%2Cincitar-a-interfer%C3%Aancia-em-processos-democr%C3%A1ticos>

<https://twitter.com/YouTubeBrasil/status/1587018223307948033>

Link to Attachment

No Attachment

2023-001-FB-UA

PC-11009

Latin America

Case number

Public comment number

Region

Thales

Bueno

English

Commenter's first name

Commenter's last name

Commenter's preferred language

London School of Economics and
Political Science - LSE

No

Organization

Response on behalf of organization

Short summary provided by the commenter

This public comment is divided in two parts and draws upon the four bullet points asked in the case. Part 1 provides a contextual analysis of Brazilian elections. How has it changed after Lula's victory? What were key mistakes committed by Meta during the period? In Part 2 I write suggestions on how Meta should review its crisis protocol, considering 3 main fields: 1) Human Rights and international Law as the bedrock of content policies 2) pré-elections risk approach 3) During elections risk approach.

This comment contain only the 2nd part of the comment.

Full Comment

Introductory remarks:

- 1) Meta did the right choice removing the content and applying penalties for the user;
- 2) The appointed case should work as a learning episode, helping the oversight board to **influence process changes at Meta's policies on a large scale, applying it for different political contexts.**
- 3) I am an **ex-vendor from Meta's Governments, Politicians and Non-Profits (GPN)**, working directly in the elections plan of 2022 and this comment is not a personal revenge, but a reflection exercise.

Part 2: Meta's election integrity time-frame and how to identify illegal political speech at a local level.

Below I present a revision over the integrity process considering 3 layers; Layer 1 is a Macro-review over content moderation, related to human rights and workforce training. Layer 2 contains recommendations on how to act before the election

period starts. Layer 3 contains suggestions related to the period of elections, including how the time-frame should work after the official result.

Macro review:

The first step is a clear definition over Meta's commitment to human rights and elections integrity that should be reflected on content moderation practices. How Meta's policies under political issues will be reflected on human rights? There's enough academic and civil society examples on how to do it and this comment will just reinforce the importance of it. That is the top-down layer that should be clear to guide integrity policies and moderation actions. Considering that the manual content moderation work happens guided by playbooks, beyond new texts, moderators should be trained in that direction. Building upon that, training a different class of moderators with focus on elections, from different languages, could be an interesting experiment.

Pre-election:

The second step is a local approach, as contextualized as possible. How do national and international human rights organizations evaluate the political context of the country under analysis? How do consultants of political risk analysis evaluate the **elections in the country? What's the importance of Meta's platform over the** country? Those questions should guide the risk analyses before the election process in each country.

Departing from the result of this evaluation, Meta should build a first time-frame of integrity efforts according to the level of risk and set up a pace of risk review during the proposed period. The key recommendation here is in the governance of the process. The risk review should be supported by trusted organizations committed to human rights and democracy oversight. In other words, Meta should not take decisions of risk analysis by itself.

During elections:

The third step considers possible changes on the integrity efforts during the election period. Based on the description provided by the case, there were two mistakes that justified the discussion of it:

- 1) The content was not escalated for Policy and Legal teams even after two appeals.
- 2) It seems that the integrity workforce underestimated the actions after January first, considering that the roots of the mobilization happened mainly over Telegram.

Suggestion 1: Less Meta, more Governance. As soon as the results of certain elections were recognized by civil society and international bodies, Meta should update their policies in the same way as Youtube did (as mentioned in my previous comment). That kind of text provides a clear guideline to route the actions of content moderators at the ground level.

Suggestion 2: Improve the guidelines related to the escalations of political issues. How many flags/appeals the post had received? What was the feedback from trusted

flaggers? Based on questions like these, content moderators should escalate the case for Meta's policy and legal teams.

Suggestion 3: The risk monitoring work made by Meta should look over other platforms and not only Instagram and Facebook. This example under discussion shows that the roots of the riots became on telegram, but were reflected on Instagram, Facebook, Twitter and WhatsApp. Even considering the closed nature of WhatsApp, a broader network analysis would show the spread of anti-democratic behaviors that would put the integrity team over alert.

Suggestion 4: Building upon the 3 previous suggestions and answering the question of time-frame, it is not possible to suggest an unique time-frame, but it is possible to suggest milestones.

Milestone 1: Calculation of the results, local and international recognition of trusted sources.

Milestone 2: Transition of power, peaceful process.

Meta should consider a band of one week after both milestones with daily check-ins backed by a cross-network analysis. Based on the risk analysis and supported by external organizations, Meta would be well informed to measure risks, demobilize resources and justify it externally.

As a final recommendation for the Oversight Board, having access to the content as it was posted would improve the level of analysis.

[Link to Attachment](#)

No Attachment

2023-001-FB-UA

PC-11010

United States and Canada

Case number

Public comment number

Region

Susan

Benesch

English

Commenter's first name

Commenter's last name

Commenter's preferred language

Dangerous Speech Project

Yes

Organization

Response on behalf of organization

Short summary provided by the commenter

Meta should convene diverse groups of experts to advise it on election related content, and should prioritize moderating speech that increases the risk of intergroup violence.

Full Comment

Electoral periods require the most vigorous protection of public peace and freedom of expression since without both, elections cannot be properly conducted, and **democracy may be undermined. Meta must act accordingly, under the Board's guidance.**

Election denialism, defined by the ACLU as "baselessly casting doubt on or refusing to accept the outcome of free, fair, and secure elections," undermines public confidence not only in a particular election or electoral process, but in public institutions. Election denialism is also routinely and directly linked to political violence since – as in this case – it is a powerful justification. It turns the tables, making political violence seem virtuous: if elections have been stolen, those who protest, even violently, can see themselves as brave, honorable guardians of democracy, a nation, and/or an identity group.

Meta must be alert to the fact that election denialism usually begins long before **votes are cast, with assertions that the relevant institutions and technology aren't adequate to conduct fair elections, and that the people involved are corrupt.** This prepares followers to believe a subsequent assertion that an election was rigged, and may discourage voting. Denialism is a way of obstructing a fair election, for example when candidates baselessly predict fraud to try and secure victory whether **they win legitimately or not, telling their followers something like: "if they tell you I lost, you will know the election was fraudulent."**

Many elections are not free, fair, and secure, however, and it is vital to allow criticism and inquiry into possible flaws and irregularities.

To make many key judgments and decisions regarding elections, Meta needs advice from external experts with specialized knowledge about elections, and electoral processes in specific countries. Well in advance of an election like the Brazilian one, when there are warning signs for election denialism and coordinated violence, Meta should convene a body of external experts to advise on country conditions and answer questions relevant to content moderation, review key moderation decisions, and flag content they consider harmful. Such bodies would work in conjunction with internal Elections Operations Centers like the one Meta assembled for the 2022 Brazilian election. They would also build on the sort of collaboration that Meta **enjoyed from Brazil’s Superior Electoral Court (TSE), since collaboration would be more extensive and would include a wider variety of participants.**

Experts might be drawn from local or national NGOs and civil society organizations, human rights defenders, scholars, journalists, an electoral administrative council and/or court like the TSE (where such an institution is highly regarded and independent), and international monitors from organizations like the Carter Center, the International Foundation for Electoral Systems, or the United Nations. The identity of experts may be kept confidential to protect them from possible **repercussions, but the expert body’s membership should be disclosed to a trusted external organization (possibly even the Oversight Board) to verify that Meta has constituted a diverse, nonpartisan body.**

Meta should consult such experts on the sort of vital questions posed here, e.g. when it is reasonable to cast doubt on an electoral process and when it is denialism, and which moderation tools are likely to be most beneficial in reducing the risk of violence. Similarly, the experts will be invaluable for distinguishing legitimate political organizing and harmful coordinated action – often the same activities and even similar language can be used for both. As an example, poll monitoring can be harmless and constructive, or intimidating to voters and poll workers.

As for the timing of Meta’s election integrity efforts, including building an external body of independent observers and civil society as described above, efforts should be launched as soon as campaigning begins in a country or region, and should continue until a peaceful transition of power has been completed and (if applicable) denialism subsides so that the transfer of power is not in doubt, nor are the relevant institutions or office holders in evident danger. Meta’s external expert bodies should advise on when these concerns have passed, and election integrity efforts can be stopped.

If the Board plans to issue a policy advisory opinion on Meta’s approach to addressing election-related disinformation and violence generally, we would also encourage review of Meta’s handling of political advertising and disparate handling of election-related disinformation and misinformation in different languages (e.g. fact-checking identical claims in English but not in Spanish).

Finally, we offer guidance regarding the Board’s last two questions. For distinguishing between legitimate political organizing and harmful coordinated action it will be invaluable to consider, with the help of external experts, what the goals and especially the likely outcomes are. Political organizing may lead to violence in fraught conditions, but it should neither be meant – nor be likely – to galvanize such violence.

The Board’s final question is arguably the most difficult, since the right to criticize government and its institutions and policies is at the core of freedom of speech. Such speech must therefore be protected by default. However as the German notion (and constitutional regime) of “militant democracy” holds, extremists must not be permitted to use democratic tools to destroy democratic institutions and processes. Meta should make use of its own existing tools to respond to content that is manifestly false and/or that has a significant chance of inciting violence. To gauge these risks regarding specific content, it should study responses to it, rely on high-quality research on the links between speech and violence, and consult experts as described above.

Link to Attachment
[PC-11010.pdf](#)

2023-001-FB-UA

PC-11011

Latin America

Case number

Public comment number

Region

Anna Luisa

Walter de

Portuguese

Commenter's first name

Commenter's last name

Commenter's preferred language

Clínica de Direitos Humanos da Pontifícia
Universidade Católica do Paraná (PUC-PR),
Brasil.

Organization

Yes

Response on behalf of organization

Short summary provided by the commenter

Os esforços de integridade eleitoral de Meta devem levar em consideração a realidade e as experiências de cada país em que possui operações. No caso brasileiro e suas eleições gerais de 2022, é notória a atuação da empresa com relativa antecedência ao processo eleitoral. No entanto, uma atuação empresarial responsável por parte de Meta em num contexto de eleições não deve se encerrar após a posse dos candidatos eleitos. A empresa deve estar atenta para quando as conturbações sociais originadas em virtude do processo eleitoral persistem após o seu encerramento formal.

Full Comment

Antes das eleições, já havia a consciência de que o período eleitoral no Brasil geraria intensos debates, com a oposição de candidatos muito populares que já haviam ocupado a presidência: Jair Bolsonaro (então Presidente) e Luís Inácio Lula da Silva.

As eleições foram marcadas por ataques feitos ao sistema eleitoral brasileiro, que geraram uma preocupação intensa do Tribunal Superior Eleitoral (TSE) sobre o uso das redes sociais para a propagação de desinformação eleitoral: as urnas, apesar de terem sido atestadas como seguras por inúmeros órgãos governamentais e agências independentes, tiveram sua credibilidade atacada constantemente pelo então Presidente e seus apoiadores durante o período eleitoral. As eleições ocorreram e culminaram na eleição de Lula no segundo turno, por uma diferença de mais de 2 milhões de votos em relação ao segundo colocado. Com a vitória de Lula, a situação política brasileira mudou. Além de insistirem na tese fantasiosa de fraude às urnas eletrônicas, apoiadores de Bolsonaro, inconformados com o resultado das eleições,

começaram a orquestrar uma tentativa de Golpe de Estado via Facebook, Instagram, Whatsapp e outras plataformas de mídia social para impedir a posse do novo Presidente democraticamente eleito: uma série de manifestações antidemocráticas foram organizadas ao redor do país clamando por um golpe militar.

A situação, que antes consistia apenas de ameaças feitas por Bolsonaro e seus apoiadores, se tornou concreta: o Brasil sofreu, no dia 8 de janeiro de 2023, uma tentativa de Golpe de Estado, orquestrada por meio de plataformas de mídias sociais.

No Brasil, há uma relação complexa entre violência política, negacionismo eleitoral e apelos à mobilização offline nas redes sociais, observada em diversos contextos políticos recentes no país, incluindo as eleições presidenciais de 2018. A violência política é preocupação crescente no Brasil nos últimos anos, com assassinatos e ataques a políticos e ativistas, inflamados pelas redes sociais, usadas para disseminar teorias conspiratórias, difamar adversários políticos e incitar o ódio e o negacionismo eleitoral.

Neste sentido, os esforços de integridade eleitoral de Meta devem levar em consideração a realidade e as experiências de cada país em que possui operações. No caso brasileiro e suas eleições gerais de 2022, é notória a atuação da empresa com relativa antecedência ao processo eleitoral, seja mediante a assinatura de um memorando de entendimentos com o TSE em fevereiro de 2022, diminuindo a frequência de conteúdos políticos no feed dos usuários três meses antes das eleições e a criação de um centro de monitoramento de conteúdos sobre o processo eleitoral. Dessa forma, entendemos que a adoção de políticas próprias mesmo antes do início do período das eleições configura um prazo razoável.

No entanto, seria ingênuo acreditar que uma atuação empresarial responsável por parte de Meta em num contexto de eleições deva se encerrar após a posse dos candidatos eleitos. A empresa deve estar atenta para quando as conturbações sociais originadas em virtude do processo eleitoral persistem após o seu encerramento formal.

Levando em consideração as publicações explícitas de “pegar em armas e forçar a entrada em prédios públicos”, chama a atenção que o conteúdo, do caso analisado, tenha permanecido no ar mesmo após a análise de 7 moderadores.

Aponta-se, portanto, três explicações plausíveis para essa imperícia:

É preciso saber qual a familiaridade destes moderadores com o contexto brasileiro, ante ao silêncio sobre a fluência dos moderadores em português do Brasil e se esse número foi incrementado durante as eleições; segundo, a atuação dos moderadores pode ter sido prejudicada pela redação da Política de Violência e Incitação, que proíbe declarações que incitem que armas sejam levadas à locais associados unicamente com o processo eleitoral propriamente dito (ex: zonas eleitorais); finalmente, destaca-se que, à contrapartida de outras plataformas digitais, a Meta não possui uma política de moderação de conteúdo própria para integridade eleitoral, e que mesmo na existência de uma, devem ser observados os reflexos

desse processo eleitoral, gerando uma agitação política, e quanto tempo essa inquietação pode perdurar, assim seria uma política eficiente.

Para além das suas próprias políticas de comunidade, a Meta também poderia usar de amparo, para realizar distinção entre o que é discurso político legítimo e atuação coordenada danosa, documentos internacionais de direitos humanos. Não se trata necessariamente de uma inovação, dado que a companhia reconhece a autoridade e se compromete a respeitar os Princípios Orientadores de Empresas e Direitos Humanos, mediante sua Política Empresarial de Direitos Humanos. Mas, no caso em voga, resta evidente que se padrões internacionalmente reconhecidos fossem levados em consideração, a postagem teria sido prontamente removida. No caso brasileiro, membro do Sistema Interamericano de Direitos Humanos, a Carta Democrática Interamericana é explícita em seu artigo 4º quanto à importância do respeito ao Estado de Direito pela sociedade para o devido funcionamento democrático, bem como a sujeição constitucional das autoridades do Estado à autoridade civil legalmente constituída. Distinções entre o exercício legal de direitos, como a liberdade de expressão e direito à reunião pacífica, e restrições que visam garantir a segurança da população e os interesses de uma sociedade democrática, são há muito reconhecidas em documentos universais de direitos humanos, como o Pacto Internacional sobre Direitos Cíveis e Políticos. Se a Meta deseja distinguir entre o exercício regular de um direito e seu exercício abusivo, e não consegue unicamente por meio de suas políticas de comunidade, certamente o direito internacional dos direitos humanos fornecerá as respostas que busca.

Link to Attachment

[PC-11011.pdf](#)

2023-001-FB-UA

PC-11013

Central and South Asia

Case number

Public comment number

Region

Noor

Waheed

English

Commenter's first name

Commenter's last name

Commenter's preferred language

Digital Rights Foundation

Yes

Organization

Response on behalf of organization

Short summary provided by the commenter

The problematic nature of the post under consideration is contingent on two factors: fake news/misinformation and political volatility. Meta has a responsibility to uphold the rights of its users to protest and express themselves and their political views on its platforms. However, given the potential for imminent violence, Meta also has the duty to preserve human life and security offline. Meta must strive to achieve a balance based on human rights standards since, historically, restrictions **under the pretext of “maintaining order” have been exploited by repressive regimes.** Meta cannot ignore that the right to peaceful protest (including protesting election outcomes) is also a fundamental characteristic of a democracy.

Full Comment

In September 2022, an op-ed published in Al-Jazeera surmised that in case of the loss of incumbent contender, former president Jair Bolsonaro, Brazil may experience something equivalent to the January 6th storming of the Capitol in the US because Bolsonaro had been making false claims of election fraud and rigging. Even without the benefit of hindsight, experts and human rights defenders on the ground were raising the alarm for a potential situation like the one that occurred on January 8th. The problematic nature of the post under consideration is contingent on two factors: fake news/misinformation and political volatility. Meta has a responsibility to uphold the rights of its users to protest and express themselves and their political views - regardless of political leaning - on its platforms. However, given the potential for imminent violence, Meta also must temper this responsibility with the duty to preserve human life and security offline by moderating posts

inciting violent protests in the interest of maintaining peace and public safety. In doing so, Meta must strive to achieve a balance based on human rights standards **since, historically, restrictions under the pretext of “maintaining order” have been** exploited by repressive regimes to curb freedom of expression. Additionally, bans on grounds of political volatility may have a detrimental effect on legitimate self-determination movements, especially in places where electoral processes are less secure and may warrant citizen protests, civil disobedience and social media speech to uphold true democratic sentiment. Meta cannot ignore the fact that the right to peaceful protest (including protesting election outcomes) is a fundamental characteristic of the democratic process.

Regarding the aforementioned posts, Meta should conduct both a risk assessment and sentiment analysis that takes into consideration (1) the potential for the outbreak of violence, (2) the volatility of the protestor, and (3) the potential for virality in the particular context. It is important to point out that if this is implemented at scale by automated systems, then guardrails should be present in the form of human reviewers from the local context who can make an analysis based on the power relations between the poster and the target of the post.

Aggressive language, in and of itself, cannot be a criterion for removing content. If the sentiment analysis shows a predominance of aggressive and violent sentiment, and if the risk assessment shows a likelihood of imminent violence, then the post should be flagged.

Furthermore, there should also be a distinction based on the position and status of the poster, i.e. in the capacity of a civilian or state actor. Posts containing calls to action issued by state actors have a higher potential for violent outcomes and, depending on the volatility of the political situation, should be taken down.

Additionally, whenever doubts are raised regarding the validity and authenticity of the electoral process and results, social media companies must work with independent fact-checkers and third-party civil society claims. The impact of unsubstantiated claims can be mitigated through a disclaimer based on factual information from official and third-party sources, in a manner similar to that deployed during the Covid-19 pandemic.

Regarding election integrity efforts, it is important to account for post-election violence, especially since, in most democracies, the transfer of power is slow, subject to an interim period, and in some cases, overseen by caretaker governments. In these contexts, specificities of the political system should be taken **into account to extend integrity efforts beyond the “official election period.** Lastly, it is essential for Meta to invest specific resources for election integrity in each context as there is no one-size-fits-all solution for the complex and diverse political systems it operates in.

Link to Attachment

[PC-11013.pdf](#)

2023-001-FB-UA

PC-11015

Latin America

Case number

Public comment number

Region

Victor

Carnevalli

Portuguese

Commenter's first name

Commenter's last name

Commenter's preferred language

Coalizão Direitos na Rede

Yes

Organization

Response on behalf of organization

Short summary provided by the commenter

O cenário político e social brasileiro apontavam explicitamente para a iminência de atos violentos, atentatórios contra o Estado Democrático de Direito, de violência contra membros de instituições de Estado e depredação de patrimônio público. O caso, portanto, demonstra a necessidade da Meta em desenvolver regras de devido processo mais robustas, nas quais é todas as decisões que impactem conteúdos de usuários devem ser justificadas e motivadas. Além disso, há que se aumentar a transparência sobre a capacitação dos times de moderadores de conteúdo. Por fim, indicamos caminhos para análise de conteúdo danoso, odiosos e desinformativos.

Full Comment

É imprescindível destacar o perfil demográfico do Brasil nas redes sociais: o Brasil tem a maior população e audiência online na América Latina, além de ter a quinta maior presença em redes sociais no mundo. Dos seus 214 milhões de habitantes, 165 milhões são usuários das redes sociais. Outro elemento torna essa presença ainda mais significativa: com uma média diária de 03h46min, é também o segundo do mundo que mais gasta tempo em redes sociais por dia; Entre outubro de 2022 e 08 de janeiro de 2023, dia em que a incitação de ataque a instituições democráticas culminou em resultados concretos, houve a intensificação de problemas que já eram correntes no país: polarização política, desinformação, discurso de ódio e ataques à imprensa e a instituições democráticas, questionamento de pesquisas, ataques a instituições democráticas e aos ocupantes de cargos públicos que os compõem - em especial, o TSE e o STF, informações falsas e enganosas sobre o processo eleitoral de modo geral;

Para evitar desdobramentos violentos, a meta precisa fazer uma análise da manifestação online, apreciando a iminência de danos, o risco de violência, quem está falando e a vulnerabilidade dos bens jurídicos ameaçados. No caso em apreço todos os componentes apontavam para o risco iminente de violência. Além disso, havia o histórico de incidentes políticos associados a campanhas de desinformação muito semelhantes, como o caso do Capitólio, apontando para possibilidades de ataques no Brasil. Os promotores desse conteúdo de violência tinham alta visibilidade e esse conteúdo reverberava, inclusive, em altos oficiais do Estado com elevado poder de mobilização. Disso tudo se conclui pela flagrante iminência de atos violentos que se demonstrava à época.

Por ter alta entrada em todos os setores sociais do país, as plataformas de mídias sociais são consideradas canais fundamentais de interlocução entre candidatos e atores políticos e suas bases eleitorais. Esses atores precisam ser considerados potenciais vetores de conteúdos violentos, odiosos e desinformativos. É preciso, portanto, que não haja tratamento diferenciado quanto à aplicação de termos de uso e políticas das plataformas entre estes e demais usuários.

Considerando a natureza de rede das mídias sociais, é impossível dissociar um único conteúdo desinformativo da rede de conteúdos semelhantes que estão circulando. É preciso considerá-lo como parte de algo interconectado em rede, inclusive com outros formatos e de plataformas externas;

No que se refere à violência política, negacionismo eleitoral e mobilização offline pela rede, é necessário que haja remoção célere de postagens que ameacem candidatos com violência física ou incitem violência contra mulheres, negros, indígenas, quilombolas e LGBTQIA+. Ainda, que sejam definidos protocolos de crise para remoção célere de conteúdo que chamem por atos violentos e abolicionistas do Estado Democrático de Direito.

É necessário que a Meta consiga analisar e compreender o contexto político e social de cada localidade em que seus serviços são oferecidos. Para isso, contar com analistas que compreendam esse contexto é crucial. A Meta deveria formar conselhos consultivos multissetoriais - com representantes de instituições públicas, da sociedade civil, da academia e do setor privado -, de diversas áreas, respeitando critérios de diversidade e que consiga analisar o contexto político eleitoral do país com vistas a possibilitar a decisão fundamentada e justificada da plataforma quanto à janela de utilização de esforços de integridade eleitoral;

É essencial que os esforços de integridade sejam realizados não apenas durante os dias de votação, mas sim durante todo o processo eleitoral até, no mínimo, a transição de poder. Ainda, deve existir a possibilidade de extensão do período caso a análise de conjuntura realizada pela plataforma e pelo conselho consultivo resulte nessa necessidade. É crucial que a empresa faça análise da realidade de cada país em que atua, buscando as melhores saídas para cada situação específica, com modulação temporal;

Considerando a realização de esforços de integridade eleitoral que são temporários, a Meta deve, com antecedência razoável - no mínimo 6 meses -, construir, em diálogo com o conselho consultivo multissetorial, as políticas internas que serão válidas durante o período de esforço eleitoral. Essas regras devem ser objetivas, acessíveis e devem refletir valores de direitos humanos e proteção da integridade democrática;

A Meta deve se comprometer a colocar a proteção da integridade eleitoral dos países em que atuam como um valor refletido e consolidado em suas políticas de moderação de conteúdo e em seus termos de uso. No caso do Brasil, por exemplo, é essencial que haja canais de diálogo e protocolos estabelecidos de crises em contato com a Justiça Eleitoral e o Ministério Público Eleitoral;

A Meta deve desenvolver regras de devido processo. Todas as decisões que interferem em conteúdos de usuários devem ser todas justificadas, motivadas e relatadas e com possibilidade de revisão por meio de mecanismo disponibilizado pela plataforma. O caso em análise demonstra a necessidade urgente de se criar procedimentos internos robustos de devido processo

É preciso ampliar a transparência no que refere às capacidades dos times de moderação de conteúdo da Meta. Os times precisam ser capazes de compreender e analisar o contexto político e social em que estão inseridos. Falta transparência sobre dados gerais, não individualizados, mas que possam mostrar que as pessoas envolvidas são habilitadas para tal: de quais nacionalidades são e que línguas possuem fluência, dentre outros;

Anúncios políticos devem ser catalogados e classificados como tal e devem estar disponíveis em bibliotecas de anúncios, com espaço em separado para anúncios e propaganda política. Nas 72 horas antes e depois do pleito é recomendável que se restrinja a circulação de anúncios políticos.

Link to Attachment

[PC-11015.pdf](#)

2023-001-FB-UA

PC-11016

Latin America

Case number

Public comment number

Region

João Victor

Archegas

Portuguese

Commenter's first name

Commenter's last name

Commenter's preferred language

Institute for Technology and
Society of Rio (ITS Rio)

Organization

Yes

Response on behalf of organization

Short summary provided by the commenter

O presente comentário público do ModeraLab a respeito do Caso 2023-001-FB-UA do Oversight Board tem por objetivo (1) apresentar alguns elementos contextuais sobre a situação do Brasil antes e após as eleições presidenciais de 2022, com um foco no papel das plataformas digitais; (2) indicar ao Board que seu processo decisório está inserido em um momento no qual o país está debatendo com mais vigor a pauta da regulação de plataformas; (3) indicar, à luz de nossos comentários públicos anteriores, sugestões de aprimoramento para a política de crise da Meta; e (4) ressaltar a importância da participação de moderadores com conhecimento local dos fatos e de uma maior fundamentação de suas decisões.

Full Comment

See attachment for full comment.

Link to Attachment

[PC-11016.pdf](#)

2023-001-FB-UA

PC-11017

United States and Canada

Case number

Public comment number

Region

Saurav

Ghosh

English

Commenter's first name

Commenter's last name

Commenter's preferred language

Campaign Legal Center

Yes

Organization

Response on behalf of organization

Short summary provided by the commenter

Based on our organization's expertise in campaign finance regulation and preservation of democratic institutions and norms, Campaign Legal Center makes specific recommendations herein on the following topics: (1) the importance of transparency as an election integrity tool; and (2) the risks of harm posed to the democratic process by mis- and disinformation disseminated during the period between when ballots are cast and election results are finalized, up to and including the assumption of office by the candidate(s) who received the most votes.

Full Comment

Please see attached file.

Link to Attachment

[PC-11017.pdf](#)

2023-001-FB-UA PC-11018 United States and Canada

Case number

Public comment number

Region

William

Adler

English

Commenter's first name

Commenter's last name

Commenter's preferred language

Center for Democracy & Technology

Yes

Organization

Response on behalf of organization

Short summary provided by the commenter

Election integrity efforts and crisis response protocols must take into account the unique context of the election, the impact election-related speech can have beyond the election period, and the impact and newsworthiness of speech by public figures.

Full Comment

The user posted on January 3, 2023, two days after the swearing in of President Luiz Inácio Lula da Silva and five days before the attack on the National Congress. At this point, supporters of the former President had already been camping outside the National Congress since his electoral loss; on December 12, supporters of former **President Bolsonaro clashed with security forces in the nation's capital. A text overlay on the video reading "we demand the source code" references the electronic voting systems used to facilitate and tally the election results. This caption evokes the narrative, widespread among Bolsonaro's supporters, that the voting system was rigged to cause Bolsonaro's electoral loss. As our report on election disinformation in Brazil, the US, and France finds, this narrative was a common element in disinformation campaigns that have circulated across Brazil since 2018.**

This context is critical to underscore the unique state of Brazilian elections and to situate the post at the heart of this case within a longer lineage of speech intended to sow distrust and **mobilize the nation's citizens to reject the results of elections. A reviewer familiar with Brazil's elections may have been able to ascertain quickly the post's intended impact. Although Meta indicated that they were staffing up their local Brazil moderation team before the elections, we do not know the expertise of the six reviewers who reviewed this post and whether they had any familiarity with Brazil and their elections.**

Part of Meta's election integrity efforts includes the Elections Operations Center, which according to Meta monitors a range of issues in real time, including voter **suppression content and other interference with individuals' ability to vote.** When and for how long this Elections Operations Center was activated remains unclear. Although we are not certain about the extent to which special efforts were in place for the Brazil election, it is likely that in Brazil election integrity efforts were rolled back soon after the vote if what the company did after the United States election is any indication. According to media reports, Meta rolled back its election integrity efforts one month after the 2022 election in the United States.

It is not entirely clear when in an election cycle it is appropriate to ramp down moderation of harmful election-related content. The aftermath of the 2020 U.S. presidential election indicated the danger of ramping down efforts to moderate election-related content in the weeks following an election, before the transition of power. The attacks on the National Congress in Brazil indicate that perhaps platforms should remain vigilant even after a transition of power. As elections **experts write in Lawfare, both the US and Brazil "would benefit from a more prolonged period of prudence by social media companies following election day (if not permanently)." Meta should be more transparent about how it determines the conditions under which its election-related efforts, such as its "Elections Operations Center," are activated and ramped down. Meta should also consider the possibility that a model whereby election-related operations are fully deactivated by an arbitrary date may leave the company unprepared to handle a sudden uptick in violence-inciting narratives and should continue to monitor developments once it can conclude any threat of violence has abated.**

In the aftermath of the U.S. attack on the Capitol in January 2021, and the Facebook Oversight Board case regarding former President Trump, Meta rolled out its Crisis Policy Protocol to address risks to individuals during temporary high-risk incidents. Whether or not the Crisis Policy Protocol was invoked in this case and what it entails remains unclear. Meta must disclose what this Protocol entails and the criteria it uses to decide when an incident is quickly changing and requires more staff to monitor.

The questions at the center of this case are not only related to whether the **company's policies are sufficient, they are also related to whether the existing suite of election and incitement of violence policies are equitably enforced.** In order to **facilitate accountability and oversight over the company's efforts and application of its policies,** the Oversight Board and global civil society organizations require greater transparency from the company, particularly around the **company's enforcement of content policies across languages and regions.**

Recommendations:

- Meta should create oversight mechanisms or checkpoints where human reviewers note down a reason/rationale behind a content-related decision, at least for posts that have been reviewed multiple times or have been appealed. This would enable

the OB to understand the thought process behind existing policy enforcement and better identify opportunities for intervention to bring enforcement decisions closer in line with the stated policies.

- In the lead up to an election, Meta should engage a host of experts including election authorities, academics, social media platforms, journalists, election officials, and civil society, to monitor and mitigate election-related speech and extend these efforts till after the certification of the vote and in quickly changing conditions till the experts deem fit.

- Meta should disclose when, where, and in what languages they deploy automated content analysis tools to enforce election-related policies. It should also disclose to the user and the Oversight Board when a post has been taken action on by an automated system.

- Meta should disclose when, for what reason, and for how long it ramps up its election integrity efforts around an election and what that entails.

- Meta should also disclose the number of reviewers in a certain language they deploy in general, in the lead up to an election, and till when.

Link to Attachment

[PC-11018.pdf](#)

2023-001-FB-UA

Case number

PC-11019

Public comment number

Latin America

Region

Ester

Commenter's first name

Borges

Commenter's last name

Portuguese

Commenter's preferred language

InternetLab

Organization

Yes

Response on behalf of organization

Short summary provided by the commenter

As eleições de 2022 evidenciaram que as plataformas digitais se concretizaram como verdadeiras infraestruturas do debate público, carregando parte significativa das conversas sobre assuntos de interesse da vida cívica. Nossa contribuição contextualiza o que InternetLab e parceiros observaram no monitoramento de um ecossistema multiplataforma de desinformação sobre integridade eleitoral. Entre nossas recomendações para a Meta no combate a esse tipo de narrativa em suas plataformas, observamos a necessidade de desmembramento de uma política de integridade cívica e democrática em dois ramos que lidariam tanto com a proteção mais ampla dos processos democráticos como, subsidiariamente, com as especificidades de momentos eleitorais.

Full Comment

Nossa contribuição está anexada.

Link to Attachment

[PC-11019.pdf](#)

2023-001-FB-UA

PC-11020

Latin America

Case number

Public comment number

Region

Victor

Carnevalli

Portuguese

Commenter's first name

Commenter's last name

Commenter's preferred language

Coalizão Direitos na Rede

Yes

Organization

Response on behalf of organization

Short summary provided by the commenter

O documento traz apontamentos e sugestões sobre o contexto brasileiro, questões afetas à transparência, escala das plataformas digitais, recomendações e conclusões. Importante destacar que o caso em questão evidencia a necessidade de desenvolvimento de regras robustas de moderação de conteúdo, com decisões justificadas e motivadas. Além disso, é necessário expandir a transparência sobre a capacitação dos times de moderação de conteúdo. Por fim, indicamos caminhos para análise e ações sobre conteúdos desinformativos, odiosos e violentos.

Full Comment

Introdução

1.1 O Brasil possui uma presença populacional massiva nas redes sociais, sendo o maior em número na América Latina e quinto no mundo. De seus 214 milhões de habitantes, mais de 165 milhões são usuários de redes sociais. Além disso, o país é o segundo do mundo que mais gasta tempo em redes sociais, em média 3h46min, de acordo com a pesquisa da plataforma Cupom Válido.

1.2 Entre outubro de 2022 e 08 de janeiro de 2023 a mudança que pode ser observada quanto aos conteúdos disseminados nas redes é que, enquanto o período eleitoral estava em aberto havia conteúdos destinados a minar a confiança no sistema eleitoral nacional, atacar a imprensa e institutos de pesquisa e propagar desinformação negativa quanto a candidatos. Após o anúncio dos resultados, os conteúdos disseminados partiram para a explicitação de insubordinação quanto aos resultados do processo eleitoral, alegação de fraude e chamamento a manifestações com o intuito de abolir o Estado Democrático de Direito, com indícios de atos

violentos contra pessoas de postos institucionais e contra prédios do patrimônio público. Tais chamamentos se concretizaram no dia 08 de janeiro de 2023.

1.3 As narrativas foram especialmente eficientes no questionamento da credibilidade das urnas eletrônicas, submetidas a testes de segurança e confiabilidade e nunca tiveram sua lisura comprometida. Ainda assim, campanhas desinformativas eram reincidentes com alto investimento em impulsionamento de conteúdo, com recapitulação de conteúdos desmentido por especialistas e anteriormente removidos das plataformas.

Questões afetas à Transparência

2.1 Falta transparência no que refere às capacidades dos times de moderação de conteúdo da Meta. Os times responsáveis pela moderação de conteúdo de uma localidade, especialmente em períodos eleitorais, devem ser capazes de compreender e analisar o contexto político e social em que estão inseridos.

Recomenda-se a publicidade de dados gerais desses times, não individualizados, que possam mostrar que as pessoas envolvidas são habilitadas para tal. Por exemplo, de quais nacionalidades são e que línguas possuem fluência, dentre outros. Além do mais, é necessário transparência na forma de seleção, tanto do time próprio quanto de serviços terceirizados, e capacitação dos selecionados;

2.2 A transparência é basilar quando se faz necessário o entendimento pela sociedade do funcionamento das plataformas. Relatórios de transparência com dados sobre moderação de conteúdo, uso de mecanismos automatizados, bases da moderação de conteúdo, dentre outros, é fundamental a uma sociedade calcada em princípios do Estado Democrático de Direito.

Escala das Plataformas Digitais

3.1 Por ter alta entrada em diversos setores sociais do país, as plataformas de mídias sociais são consideradas canais relevantes de comunicação entre candidatos e suas bases eleitorais. Por esse motivo, atores políticos e ocupantes de cargos públicos são potenciais vetores de desinformação, discurso de ódio e violência política. Não cabe, portanto, às plataformas digitais oferecerem tratamento diferenciado a esses atores, colocando-os em um patamar superior ao dos demais usuários.

3.2 Sobre o caso em questão, é necessário compreendê-lo como parte de algo maior. Campanhas desinformativas e de chamamento amplo para manifestações offline devem ser compreendidas como um ecossistema de conteúdos semelhantes que, mesmo não idênticos, apontam para um mesmo norte, inclusive em outros formatos e de plataformas externas.

3.3 No que se refere à violência política, negacionismo eleitoral e mobilização offline pela rede, é necessário que haja remoção célere de postagens que ameacem candidatos com violência física ou incitem violência contra mulheres, negros, indígenas, quilombolas e LGBTQIA+. Ainda, que sejam definidos protocolos de crise para remoção célere de conteúdo que chamem por atos violentos e abolicionistas do Estado Democrático de Direito.

Recomendações

4.1 Meta deve ouvir especialistas multissetoriais, respeitando critérios de diversidade, que consigam analisar o contexto político eleitoral do país com vistas a embasar a decisão da plataforma quanto à janela de utilização de esforços de integridade eleitoral

4.2 É essencial que os esforços de integridade sejam realizados não apenas durante os dias de votação, mas sim durante todo o processo eleitoral até, no mínimo, a transição de poder. Ainda, é necessário que exista a possibilidade de extensão do período caso a análise de conjuntura realizada pela plataforma, apoiada por um corpo de especialistas, resulte nessa necessidade

Conclusões

5.1 Considerando a temporalidade de esforços pela integridade eleitoral de localidades diversas, é preciso que a Meta, com antecedência de, no mínimo 6 meses, desenvolva e publique quais serão as regras que estarão em vigor durante o período. Estas devem ser objetivas, acessíveis e devem refletir valores dos direitos humanos e da proteção do Estado Democrático de Direito. Essas regras precisam ser específicas para cada país, levando em as realidades sociais e institucionais;

5.2 De forma geral, é recomendável que a Meta estabeleça regras gerais e constantes de devido processo que consolidem o procedimento de moderação de conteúdo e os direitos dos usuários em recorrer dessas decisões, possibilitando assim canais de contestação acessíveis e eficientes. Todas as decisões que impactam na circulação de conteúdo publicado por usuário devem ser justificadas, motivadas e relatadas. O fato de seis moderadores terem tomado decisões injustificadas demonstra a necessidade urgente de consolidação de regras de devido processo.

5.3 A Meta deve se comprometer a colocar a proteção da integridade eleitoral dos países como um valor refletido em suas políticas e termos de uso, levando em conta a realidade legal, política, social e econômica da localidade e sua a realidade institucional. No caso do Brasil, por exemplo, é essencial que haja canais com Justiça Eleitoral e Ministério Público.

Link to Attachment

[PC-11020.pdf](#)

2023-001-FB-UA

PC-11021

Latin America

Case number

Public comment number

Region

Rafaela

Silva

Portuguese

Commenter's first name

Commenter's last name

Commenter's preferred language

Instituto de Referência em Internet e
Sociedade (IRIS)

Yes

Organization

Response on behalf of organization

Short summary provided by the commenter

O Instituto de Referências em Internet e Sociedade (IRIS) buscou discorrer sobre a situação política do Brasil, a partir das eleições de 2022 e o episódio antidemocrático do 08 de janeiro de 2023, trazendo reflexões sobre o tratamento de conteúdo que deslegitima instituições e processos democráticos, apresentando caminhos possíveis para ações do Meta que visem a manutenção da democracia em suas plataformas digitais.

Full Comment

Entende-se que as ações da Meta devem ser direcionadas a barrar violências políticas como uma diretriz de caráter permanente, dedicando atenção maior para cenários nacionais que estão em crise política, com parâmetros pré-definidos para determinar a existência dessas crises e atuar de forma ainda mais cuidadosa na moderação de conteúdo nesses locais. Contudo, as políticas específicas para o período eleitoral devem existir, de maneira complementar, considerando sua excepcionalidade, independente da existência de crises concretas.

De início, é importante destacar os grandes riscos existentes na circulação de conteúdos que atacam e deslegitimam instituições democráticas em plataformas digitais. Deve ser explicitamente proibida a disseminação e publicação de conteúdos com temáticas golpistas, autoritárias e/ou incitando violências contra a democracia, diante da alta capacidade de influência das plataformas geridas pela Meta.

Paralelamente, para uma moderação de conteúdo democrática, nos posicionamos pela maior transparência nos modos como essas ações são realizadas e o fortalecimento das medidas de devido processo, aliadas à construção de

mecanismos para resposta rápida e concreta da empresa, como já reconhecido pelo Oversight Board na decisão sobre a suspensão de conteúdo publicado pelo ex-presidente Trump, relacionado à invasão do Capitólio em 6 de janeiro de 2021. Assim, será possível um controle social eficaz, que identifique as práticas utilizadas pela plataforma, suas consequências, além de possibilitar a sugestão de eventuais aprimoramentos.

Em face disso, a avaliação do contexto é crucial para definição das medidas adequadas a serem tomadas, sobretudo quando envolvem indivíduos influentes, como líderes políticos. Nesse último caso, recomenda-se a garantia da revisão humana de decisões automatizadas e o encaminhamento imediato a uma equipe autônoma e imparcial, com conhecimento sobre o cenário político e linguístico e treinamento técnico, capaz de inferir as nuances da situação analisada e decidir com maior acurácia.

As normas que gerem a moderação de conteúdo devem ser explícitas e abertas para todos usuários, principalmente em se tratando de questões eleitorais. Sugere-se que a pessoa usuária seja informada, pelo menos: a) sobre a fundamentação da decisão tomada, com distinção entre avaliações políticas legítimas, atos perigosos e atos ilegais, especificando-se as violações ocorridas; b) prazo para contestar e o meio pelo qual fazê-lo, assim como o prazo para reavaliação pela plataforma; c) se a decisão foi automatizada ou não; d) a penalidade específica aplicada ao conteúdo, assim como quanto a seu caráter definitivo ou temporário, com o período de suspensão.

Diante de situações com elementos imprevisíveis, espera-se, igualmente, que a plataforma possua parâmetros de conduta para lidar com casos críticos e, quando envolverem perigo de dano iminente, subsiste o dever de resposta rápida pela empresa. Dessa forma, defende-se a projeção de um protocolo, que informe os modos com que se lidará em situações políticas, em contextos distintos, com normas que devem ser seguidas por todos candidatos numa eleição.

Recomenda-se, ainda, que as sanções de exclusão definitiva sejam utilizadas como último recurso, de modo que outras penalidades, tais como a suspensão temporária por tempo determinado (passível de renovação), a inserção de links de direcionamento para portais oficiais e restrições noticiadas ao alcance das publicações, sejam consideradas vias prioritárias, sobretudo se não houver conclusão da apuração de violação.

Por fim, as medidas de enfrentamento devem considerar as particularidades culturais e tecnológicas dos diferentes modos de disseminação de conteúdo que ataque ou deslegitime instituições democráticas. A título exemplificativo, o enfrentamento a conteúdo antidemocrático não deve ensejar medidas que fragilizem a criptografia nos serviços de mensageria oferecidos dentro das plataformas do Meta, uma vez que é um mecanismo de proteção à privacidade e dados pessoais que garante o exercício e manutenção dos direitos no ambiente digital. Nesse contexto, acreditamos que, a partir de uma moderação dedicada à

restrição de conteúdos antidemocráticos e adequadamente procedimentalizada para fins de devido processo legal, será possível a garantia de um espaço virtual melhor.

Ainda, registre-se que o motivo pelo qual os atos de 8 de janeiro são caracterizados dessa forma diz respeito ao fato de haver uma explícita oposição ao resultado das eleições de outubro de 2022. Uma vez que tais foram realizadas de maneira legítima e em conformidade com o sistema eleitoral brasileiro, não há motivo pertinente que justifique posicionamentos contrários ao resultado apurado. Ademais, os atos apresentam caráter terrorista, vândalo e criminoso uma vez que foram realizados através da destruição desmedida e violenta do patrimônio público e contra instituições democráticas.

Dito isso, acreditamos que os atos de 8 de janeiro foram explicitamente antidemocráticos e perigosos. Assim, para que a moderação de conteúdo da Meta também tenha essa compreensão, é necessário que o contexto brasileiro seja considerado, bem como quaisquer contextos em que a empresa atua. Desse modo, é preciso que haja um regulamento prévio e práticas que conciliem da melhor forma o uso de moderação automatizada e humana, destinando atenção ao seu aprimoramento também quando isolados. Torna-se importante que os conteúdos sejam analisados por indivíduos nativos ou com conhecimentos específicos sobre o país em questão, além do treinamento adequado da IA.

Finalmente, cabe mencionar que o ato de remover conteúdos ilícitos, caso realizado de maneira isolada, pode ser pouco efetivo. É preciso, também, que a plataforma explique o motivo pelo qual o conteúdo foi removido, aos usuários gerais e ao autor.

Link to Attachment

[PC-11021](#)

2023-001-FB-UA

PC-11022

Latin America

Case number

Public comment number

Region

Artur

Lima

English

Commenter's first name

Commenter's last name

Commenter's preferred language

Information Society Project, Yale Law School

Organization

No

Response on behalf of organization

Short summary provided by the commenter

When violence erupted in the capital, and the nation held its breath and feared the worst for democracy in Brazil, many blamed social media. Almost three months from January 8th, Meta has told us very little about its actions in advance of the attempted insurrection. In fact, we can't say for certain even what the policy was and when. As criticized as Meta's involvement and response to the Jan. 6 2021 insurrection in Washington, D.C., was, the reaction and publicly-available information on the events in Brasília don't compare. Meta owes more to Brazilians and to the citizens of embattled democracies where it operates around the world.

Full Comment

Attachment contains comment in accordance to 5-page limit specified in the call for public comments.

Link to Attachment

[PC-11022](#)

2023-001-FB-

PC-XXXXX

Latin America

Case number

Public comment number

Region

Commenter's first name

Commenter's last name

Commenter's preferred language

Red en Defensa de los Derechos Digitales

Yes

Organization

Response on behalf of organization

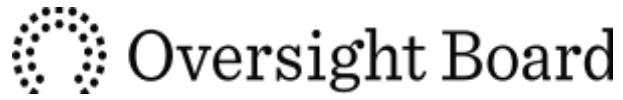
Due to a technical issue, this submission was not received in the usual manner, but it was received within the appropriate timeframe, and thus it is included in the listing of Public Comments received by the Oversight Board for this case.

Full Comment

Please See Attachment.

Link to Attachment

[PC-R3D-2023-001](#)



Public Comment Appendix for
2023-001-FB-UA

Case number

End of public comments