



Who we are

SMEX is a non-profit that advocates for and advances human rights in digital spaces across West Asia and North Africa (WANA). Our vision is for everyone living in West Asia-North Africa and the diaspora to be able to access and engage with the internet, mobile services, and other networked spaces safely and without fear of censorship, surveillance, or repercussion.

Remarks on the announcement of the case

While we welcome the Oversight Board's decision to look into the implementation of Meta's policies regarding permanently disabling accounts for targeting public figures, we notice a lack of essential information about the case at hand. The details shared are limited and fail to provide a comprehensive overview of the facts in order to assess if the disabling was rightful and in line with Meta's policies.

Firstly, unlike other recent OSB calls for submissions such as the "[Passport case](#)" this case is missing key facts. Without concrete information about the actual content of the posts that led to the permanent disabling of the account, it is difficult to evaluate if Meta's Community Standards were in fact violated and if the ban was appropriate under international human rights standards.

Again, contrary to other cases such as the "[Emojis Targeting Black People](#)" announcement, the country from which the account originated is not mentioned. This omission is particularly important since the context of expression should be taken into account, as per international human rights standards, such as the [Rabat Plan of Action](#).

The announcement doesn't mention if the posts in question were initially reviewed after being detected by Meta's human moderators or automated systems or after they were reported by a user or a trusted partner or in compliance with a government request. For transparency purposes, sharing this information is crucial.

There is also a lack of reference to the notification that the user received (if any) about their account being disabled. Not informing the users properly about any action taken against them is a violation of [international human rights standards](#), which require that such notifications contain not only a statement of reasons regarding the decision, but also an accessible process to challenge it.

How best to ensure due process and fairness to people whose accounts are penalized or permanently disabled

Any decision involving account and content restrictions should respect the principles of legality, legitimate aim, and necessity and proportionality. The Oversight Board has often relied on the “three-part” test in its [decisions](#), highlighting that any limitation on freedom of expression must be in accordance with [Article 19 paragraph 3](#) of the International Covenant on Civil and Political Rights.

The principle of legality requires that the rules imposing a restriction on freedom of expression must be clear, precise and accessible. The Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression ([A/HRC/38/35](#), para. 46) further mentions that companies should go beyond explaining their rules in detail by including aggregated data and specific examples that demonstrate how these rules are implemented. In the case at hand, not only such information is not shared, but also the policies under which the account was banned do not fulfill the requirements of clarity and precision.

International human rights standards, such as the [UN Guiding Principles](#) on Business and Human Rights, also require that human rights and due process are taken into consideration and integrated by companies throughout the process of content moderation. The [Santa Clara Principles](#) on Fairness and Accountability in Content Moderation, which Meta endorsed in 2018, also provide operational guidelines for how to do this. To ensure due process and avoid discriminatory practices, account removals should be standardized to the greatest extent possible. This also necessitates clear and comprehensive information about the removal process, the criteria used to reach such a decision, as well as detailed explanations linking to the policies allegedly violated. Meta needs to fully implement the [multiple recommendations](#) made by the Board regarding the enforcement of strikes, such as in the case of [Reporting on Somaliland Current Affairs](#). The user should receive promptly a notification containing all these details, accompanied by intelligible instructions on how to appeal the decision.

The principle of legitimate aim requires restrictions to serve one of the aims listed in Article 19(3). Protecting the rights of others is a legitimate aim and Meta should continue taking threats of violence seriously because these threats impinge on the rights of others. However, as per [UN General Comment No. 34](#), para. 11, preventing offense is not a legitimate aim. Suspending an account should be a final step, taken only if there is actual risk of real-life harm, and not if the content is merely offensive.

For a measure restricting freedom of expression to be in compliance with international human rights law, it also needs to be necessary and proportionate to the aim pursued. This means that the less intrusive measure should be adopted, for example suspending certain features of an account rather than permanently disabling it, which should be the last measure, with regard to the severity of the policy violation.

Besides, any removal of content or disabling of accounts under an exception should also be clearly stated and explained within the relevant policy. Meta's dedicated section to [restricting accounts](#) lacks clarity and precision and does not indicate in fullness under which specific circumstances an account may be disabled. It also fails to link to policies the violation of which may lead to suspending certain features or permanently disabling an account. For example, the [Violence and incitement policy](#), referenced in this case, does not link to the restricting accounts section and vice versa.

Lastly, users should be able to appeal any actioning decision against their content and accounts. The Oversight Board has substantially commented on Meta's need to revise its appeal processes on a number of cases, including [Reporting on Somaliland Current Affairs Bundle](#), [PAO on the Removal of COVID-19 Misinformation](#) and [Iran protest slogan](#). [UN A/74/486: Report on online hate speech](#) also urges companies to put in place an accessible and transparent appeal process. Operational Principle 3 of the [Santa Clara Principles](#) provides clear guidance on how companies can provide users of an adequate and accessible appeal process, as well as clear and comprehensive details about the timeline of the appeal process, users' ability to provide supporting information, and extensive explanation of the reasons for the decision. Operational Principles also include involving human moderators outside those reaching the initial decision who are in a position to understand the language and cultural context of the content relevant to the appeal.

The information on Meta's Get back into your Meta account if it's disabled or restricted [page](#) is very vague and lacks explicit mention of whether the message informing the user also contains the reason for disabling, whether the user will definitely have the possibility to appeal ("*you may be able to ask us to review the decision*") and how, as well as concrete information about the "*specified time*" for appealing. Additionally, if the user fails to appeal within the "*specified time*", they will "*no longer be able to request a review of their account*". This is a disproportionately harsh measure that essentially deprives the user of their right to have a decision reviewed. In certain circumstances, a warning about the actioning decision may not be issued ("*in some cases, we may not issue a warning before restricting your account*"). The policy gives no specific examples of what these cases might be nor an explanation behind this diversion from international human rights law that demands that a reasoning accompanies any decision affecting an individual's fundamental rights (Articles 9,13 and 14 [ICCPR](#), Articles 5, 6 and 10 [European Convention on Human Rights](#)).

Meta goes as far as introducing a de facto "no-appeal practice" for severe violations ("*We also can't restore accounts that were disabled for severe violations.*"). By refusing to restore accounts for such violations, while stripping users of their right to have the decision about the ban on their account reviewed, the policy undermines the very essence of the appeal process and violates the principle of proportionality, as well as the right to access to remedy, guaranteed in the [UNGPs](#).

The effectiveness of measures used by social media platforms to protect public figures and journalists from accounts engaged in repeated abuse and threats of violence, in particular against women in the public eye

In its [2024 human rights report](#), Meta claimed to have improved the speed at which they respond to content reported through the Trusted Partner program, of which SMEX is a member. These trusted partners provide essential insight into local context and linguistic nuances, and report specific cases to Meta for resolution. Such engagement is vital, since trusted partners highlight structural content moderation shortcomings while addressing individual human rights violations through case work. However, insights from [SMEX's Helpdesk](#) suggest that Meta is shifting away from its commitment to maintaining a meaningful relationship and open channels of communication with CSOs. In 2025, Meta did not reply to a number of escalations made by our Helpdesk, and some accounts have remained indefinitely suspended. In the following two examples shared by our Helpdesk, political expression from minority groups is seemingly leading to improper account suspension, while accounts inciting violence in the same context are often not removed:

1. The account of an LGBTQ activist in the WANA region was temporarily suspended for having shared an image of a veiled yet braless woman, conveying the message that women have control/governance of their bodies. Although an exception of artistic content should have been made, the platform proceeded to suspend the account.
2. In the ongoing crisis in Syria, where violence and systematic targeting of women belonging to minority groups (Druze, Alawite, Kurdish) prevail, Meta has repeatedly suspended accounts of activists and women reporting on the violence, while failing to remove accounts of malicious actors spreading hate speech and inciting violence against these communities.

Based on the tickets received throughout 2025, SMEX's Helpdesk observed certain patterns in relation to account suspension. 74% of all received tickets were related to content moderation, suspended accounts, and suspended activities. Instagram remains the main platform for relevant enforcement actions, making up over 53% of all reports, followed by Facebook with 22%. The vast majority of these cases relate to account suspensions indicating that Meta platforms dominate when it comes to account disruptions. Multiple users who have gotten their accounts suspended have reported inability to specifically access their "Account status" section in Instagram settings where you can check "personal content" that has been flagged or removed, and other restrictions imposed, putting them in a position where they do not know what the issue is and by extent not being able to appeal the decision.

In addition, Meta's latest policy changes, announced by the company's CEO in [a video](#) in January 2025, put an end to independent and [IFCN-certified](#) third party fact-checkers in the United States, which were replaced by a Community Notes model, similar to the one used by the [platform X](#). While this development is currently implemented only in the US, it is [a serious setback](#) in guaranteeing the integrity of content moderation practices. The changes also included relaxing restrictions on the discussion of subjects, such as [gender and immigration](#), which are part of political debate. Under the pretext of enhancing freedom of expression on its platforms,

Meta increased the risk of exclusion and put in danger vulnerable groups such as migrants and the LGBTQ community. Lowering the standards will only exacerbate the dissemination of hate speech and incitement to violence towards these groups. [Multiple examples](#) of crimes against women TikTokers and online influencers from the LGBTQ+ community in Iraq and other countries in the [WANA region](#) demonstrate that content moderation practices of big social media platforms have proven insufficient for the protection of public figures.

Challenges in identifying and considering off-platform context when assessing threats against public figures and journalists

Social media platforms tend to invest more resources in the US and the EU and [neglect the Global Majority](#). This practice has led to serious shortcomings when it comes to good content moderation practices and therefore to gross violations of the rights to freedom of expression and access to information for populations of [different regions](#).

Understanding the linguistic nuances and the local context plays an important role in the accurate and fair implementation of content moderation policies. Previous cases have shown that without familiarity with cultural notions, misinterpretation of terms by human moderators or automated systems is prone to result in severe infringements of human rights. A stark example was Meta's blanket ban on the word "shaheed" under the [Dangerous Organizations and Individuals Policy](#). The company's understanding of the use of the word in Arabic was extremely limited, leading to widespread improper content removals. [Our submission](#) to the Oversight Board highlighted, among other things, the importance of taking into account the different meanings of the term.

In a similar case, when moderating the slogan "[From the river to the sea](#)", Meta failed to use its extensive capacity, networks, resources, and tools to conduct research into both the good faith use and misconceptions surrounding it. Instead, the Oversight Board turned to civil society for public comments and SMEX also contributed with [a submission](#). Although we always appreciate the Board's efforts to include CSOs perspectives, in this particular case, given the timing of the consultation (amid Israel's genocide in Gaza), as well as the existence of large-scale available research on the topic, Meta should have investigated it on its own initiative and by taking advantage of its substantial resources.

Another case that showed Meta's lack of investment in correctly assessing local context was [Posts sharing speeches in the Syrian conflict](#), for which SMEX also submitted [a public comment](#). The removal of the two posts under the DOI and Violence and Incitement policies was yet another striking example of Meta's inefficiency to comprehend local context and linguistic nuances. For instance, as we noted in the submission, incitement to violence against Druze and other minority communities can be coded. At the same time, Meta failed to [respect the newsworthiness](#) of this case. In our submission, we referenced a 2022 [report](#) by Business for Social Responsibility (BSR), commissioned by Meta in response to a Board recommendation, that clearly states that Meta [overmoderates content in Arabic](#).

Good industry practices in transparency reporting on account enforcement decisions and related appeals

Meta has endorsed the Santa Clara Principles, a set of recommendations addressed to industries and which comports with human rights requirements and civil society requests, aiming at improving transparency and accountability in content moderation. [The company also claims](#) to use the [UN Guiding Principles](#) on Business and Human Rights, the [International Bill of Rights](#), as well as its internal [Corporate Human Rights Policy](#) as guides for Meta teams to ensure that human rights are respected. Such commitment is often viewed as symbolic but Meta should make certain that human rights safeguards are indeed incorporated in their content moderation decisions.

We encourage the Board to take into consideration the following recommendations:

1. **Clarify information about disabling accounts in the Transparency Center:** Any policy that could lead to a severe strike and immediate and permanent disabling of accounts should contain a sentence explaining this, differentiating between temporarily suspending certain account features and permanently disabling the account in its entirety, and should link to the Transparency Center article on “Get back into your Meta account if it is disabled or restricted.”
2. **Notification about any decision imposing restrictions in all cases:** Users should always receive a notification informing them of any decision restricting their ability to engage with Meta platforms, i.e. disabling or restricting accounts, and explaining the reasons for such decision and how to appeal.
3. **Guaranteed right to access to remedy:** Every user should always have the right to request at least one meaningful human review appeal, regardless of the violation.
4. **Different human reviewers to assess appeals:** The Board should encourage Meta to implement in full its recommendations in the [“Reporting on Somaliland Current Affairs” case](#) regarding Meta’s revision of its appeal processes in order to prohibit the same human reviewers from assessing appeals against their own decisions, including on page unpublication.
5. **Implementation of recommendations for companies from the [UN Report on online hate speech](#):** To uphold its international human rights obligations, Meta should thoroughly incorporate the recommendations from the UN Report on online hate speech, in particular those relevant to publishing human rights impact assessments and ensuring contextual analysis.