sflc.in

To Meta, Inc.

Subject: Joint Representation regarding content moderation practices, account restrictions, and transparency across Meta Platforms.

Respected Sir/Madam,

We the undersigned, including civil society organisations and representatives of users in India write to you to express our concerns regarding the operation of Meta's platforms that have raised significant implications for users in India and their ability to access digital communication services and exercise their rights in accordance with Meta's publicly stated Community Standards and policies.

India represents one of Meta's largest and most significant user markets globally. Millions of people across the country rely on Facebook, Instagram and WhatsApp for communication, access to information, dissemination of news and ideas, community organising, civic participation, preservation of evidence, and professional activity. For content creators, journalists, civil society organisations, small businesses, and other users, access to these platforms also form part of their profession and livelihood. Meta's content moderation and enforcement practices therefore have a substantial impact on a significant segment of its global user base, affecting not only users' ability to communicate and access information but also their professional activities, livelihoods, and participation in public life.

Journalists and news outlets are also facing systemic suppression due to over-automated content moderation systems that lack context for investigative reporting, public interest documentation, and critical commentary. By conflating legitimate reporting on sensitive or controversial issues with policy violations, Meta's decision to comply with all take down notices, without due process, effectively silences public-interest journalism, restricting reach or outright removing vital news coverage without human oversight. This opaque suppression undermines press freedom, as well as Indian citizens' right to reliable information, and leaves journalists without timely recourse to contest improper takedowns.

Over the past few months, we have come across multiple counts of **automated removal of content on Meta Platforms** with no explanation provided to users, including no mention of the particular violation of Meta Community Standards. There have been recurring reports from users highlighting instances of shadowbanning on Instagram. This has resulted in reduced visibility of posts and accounts, unexplained reductions in content reach, and restrictions on direct messaging functionality. On WhatsApp, we have noted the **suspension or disabling of WhatsApp groups** without adequate notice or meaningful explanation. We have recorded a few such instances, and the details of the same can be found in the table **appended** to this letter.

Meta itself has publicly acknowledged instances in which automated systems have incorrectly restricted legitimate content. For instance, a video posted by the Prime Minister of India, Narendra Modi on Facebook and an Instagram Reel posted by Leader of Opposition Rahul Gandhi was temporarily restricted by an automated system. These contents were subsequently restored on the platform. While public figures and large organisations may have greater access to channels through which such errors can be brought to

Meta's attention, ordinary users generally do not have comparable access to human representatives or effective escalation mechanisms. Many users may also not be fully aware of the reporting, grievance or appeal mechanisms available to them, particularly as Meta's platform interfaces and user journeys change frequently and there is limited, accessible information to guide users on how to identify, report and challenge an enforcement action. This creates an additional barrier to obtaining meaningful redress, particularly for users who may not have the technical familiarity or resources to navigate changing platform interfaces and processes. For users whose livelihoods, journalism, civic participation, or ability to communicate depend upon these platforms, an erroneous restriction can have consequences even where the content is ultimately restored.

Furthermore, where content is wrongfully removed or its reach is significantly restricted, restoration alone may not remedy the harm caused. Content that has been unavailable or suppressed during a time-sensitive news cycle, public debate or other event may lose accumulated engagement, including likes, comments and shares, as well as the opportunity to reach its intended audience. We therefore believe that mechanisms for restoration should also address, where technically feasible, the consequences of wrongful removal or suppression, including lost reach and distribution.

According to publicly available information, many of the affected posts were restricted pursuant to "automated systems implemented under legal requirements". In such cases, it is unclear whether the restriction stems from a legal request or whether Meta has restricted the content *suo motu* using automated systems as part of its compliance with due diligence requirements under the Information Technology Rules, 2021. Upon analysing the Information Technology Act, 2000 and the applicable rules, we also note that confidentiality obligations expressly apply to blocking orders issued under Section 69A of the Information Technology Act, 2000 read with Rule 16 of the Information Technology (Procedure and Safeguards for Blocking for Access of Information by Public) Rules, 2009. There is no equivalent statutory confidentiality requirement applicable to actions taken under Section 79(3)(b) of the Information Technology Act, 2000 read with Rule 3(1)(b) of the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021. Users are often not provided meaningful information about the legal or factual basis of the restriction, the specific policy or legal provision relied upon, whether an automated system was involved, or what steps they can take to challenge the decision. Such opacity makes it difficult to distinguish between government-mandated restrictions, enforcement of Meta's own policies, and erroneous or disproportionate automated moderation. This is particularly concerning where affected content relates to political expression, satire, criticism of public authorities, journalism, or peaceful protest, all of which are central to democratic discourse.

Content moderation decisions, particularly those made or assisted by automated systems, require a basic level of transparency and accountability towards users. In circumstances where content is removed, visibility is restricted, or accounts and groups are suspended, users should be provided with sufficiently clear reasons for the action taken, information regarding the policy allegedly violated, and access to an effective and timely review mechanism.

In light of the above, we request the following:

1. **Transparency regarding the basis for content restrictions:** Please clarify how users are informed whether content has been removed or restricted pursuant to (i) a government order or a court order; (ii) enforcement of Meta's Community Standards; or (iii) automated systems implemented pursuant to legal requirements. Users should be able to understand the basis and source of an enforcement action, the specific legal authority relied upon, whether automated systems were involved, and the appropriate mechanism available to challenge the action.

   Where access to an account, content, feature, or service is **temporarily restricted** pursuant to a legal requirement, users should also be informed of the nature and scope of the restriction, the expected duration of the restriction, whether it will automatically expire or is dependent on further action by the relevant authority or Meta, and the mechanism available to challenge or seek review of the restriction. Where disclosure of the underlying legal request or order is legally prohibited, Meta should nevertheless provide the maximum information legally permissible regarding the nature, duration and review mechanism applicable to the restriction.

   Meta should adopt a standardised notice that provides this information in a clear and accessible manner, reasonable efforts should be made at providing translations especially when the content affected is in regional language.

2. **Transparency regarding legal requests:** Meta should provide greater transparency regarding government and court requests for content restrictions, including information about the nature and legal basis of such requests, the content or accounts affected, and the action taken by Meta. Where Meta is legally permitted to do so (such as actions under Section 79(3)(b) of IT Act), users affected by such requests should be informed of the basis of the restriction and the available avenues for challenge. Notifications concerning content restrictions should, to the extent practicable, identify the specific portion of content that has been flagged, including the relevant line, image, segment or timestamp where applicable; the specific Community Standards provision or local law allegedly breached; and whether the action was triggered by an automated system, a user report, or a direct government or legal request. Where content or an account is restricted pursuant to a request from an authority, journalists and media organisations should, to the fullest extent legally permissible, be informed of the source, legal basis and scope of the request.

3. **Automated Moderation Systems:** Meta should provide greater transparency regarding the role of automated systems in content moderation across its platforms and the safeguards in place to prevent erroneous, inconsistent, or disproportionate enforcement. Automated systems should not be the sole basis for consequential enforcement actions without appropriate safeguards and meaningful human oversight, particularly where the action concerns journalism, political expression, public-interest information, or peaceful protest. Meta should also assess whether its automated systems adequately account for linguistic, cultural, regional, gender and other forms of diversity among its users, particularly in India, where content is produced in numerous languages and within diverse social and cultural contexts.

4. **Protection of journalism, public-interest information and vulnerable users:** Meta should adopt moderation practices that appropriately account for the context in which content is produced and shared, particularly where content concerns investigative journalism, public-interest documentation, reporting on protests or sensitive events, or critical commentary. Legitimate reporting should not be treated as equivalent to prohibited content merely because it depicts or discusses sensitive material. Meta should also regularly evaluate the impact of its content moderation practices on freedom of expression and other rights, particularly for journalists, media organisations, vulnerable and marginalised communities, and users whose ability to participate in public discourse or access essential information may be disproportionately affected by moderation decisions.

5. **Protection of vulnerable and marginalised communities:** Meta should reverse the January 2025 changes to its "Hateful Conduct" Community Standards that reduced protections against hate speech targeting women, transgender and queer persons. Meta should ensure that its Community Standards provide meaningful protection against hate speech and discrimination affecting vulnerable and marginalised communities.Meta should also regularly evaluate the impact of its content moderation practices on the freedom of expression and other rights of vulnerable and marginalised users, including their freedoms relating to their bodies, sexualities, livelihoods, religious practices, access to information and political participation.

6. **Reasons and notice:** Where content is removed, restricted, delinked from discovery, or an account or group is suspended, please clarify what notice is provided to users, including the specific action taken and the relevant provision of Meta's policies allegedly violated.

7. **Review and appeals:** Please provide details of the mechanisms available to users to challenge moderation decisions, including when human review is available, how it can be requested, and the expected timelines for review. Users should have a meaningful opportunity to obtain human review, particularly where the restriction concerns political expression, journalism, public-interest information, or peaceful protest.

8. **WhatsApp groups:** Please clarify the circumstances in which WhatsApp accounts of individuals and groups may be suspended or disabled, the notice and safeguards available to group administrators and participants, and the mechanisms for challenging or appealing such decisions. Administrators should have an accessible mechanism to challenge such decisions and obtain timely human review.

9. **Erroneous enforcement:** Please explain how Meta identifies and remedies erroneous automated enforcement, including the process for restoring incorrectly removed content, restricted accounts, or suspended groups.

10. **Visibility, reach and distribution restrictions:** Meta should clearly recognise and address significant reductions in the visibility, reach, distribution, recommendation, or discoverability of

content or accounts within its policies and Community Standards. Users should be informed where such restrictions materially affect the distribution of their content and should receive clear reasons for the restriction and an avenue to challenge it. Where Meta determines that content or an account has been wrongfully subjected to such a restriction, it should take reasonable steps to remedy the resulting loss of reach and distribution, rather than treating restoration of the underlying content alone as sufficient redress.

We welcome an appropriate response addressing the issues raised in this representation. We would also welcome the opportunity to engage further on measures that enhance transparency, user safeguards, and accountability across Meta's products and services.

Yours faithfully,

Organisations

1. Software Freedom Law Centre, India (SFLC.in)
2. Free Speech Collective
3. Index on Censorship

Individuals

1. **Mohammad Areeb Hussain Shola**
2. **SYED Irfan**
3. **Evan Anton Arun**

sflc.in

Instances of Content Takedown and Censorship faced by Signatories

S.No	Account Owner/ Account Handle	Platform	Description of Content	Type of Action as per Notification/Email	Link
1.	SFLC.in	Instagram	SFLC.in on 17th September 2026 had published a post highlighting the takedowns faced by Alt News in relation to its reporting on the *Chauhaan* film teaser. 3 out of 5 slides of the post were restricted in India, these slides had references to “Burhan Wani”. SFLC.in shared Scroll’s report on the takedown through an Instagram Story on its public handle; the story was also taken down by Instagram through the use of automated systems. The post later became available on 25th September 2026	Restricted access through an automated system implemented pursuant to legal requirements under India’s Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules"	https://www.instagram.com/p/DdZEFu4E0UF/?stkn=MWJwZWlwNGN0cXh0Nw==

sflc.in

sflc.in