sflc.in

To X Corp,

Subject: Joint Representation regarding content moderation practices, account restrictions, and transparency on X.

Respected Sir/Madam,

We the undersigned write to you to express our concerns regarding several issues relating to the operation of X's platforms that have raised significant implications for users' ability to access digital communication services and exercise their rights in accordance with X's publicly stated Community Standards and policies. As one of the largest providers of social media and messaging services, X's content moderation practices have a substantial impact on individuals, communities, content creators, journalists, civil society organisations, and other stakeholders who rely on these platforms for communication, dissemination of information, preservation of evidence, community organising and civic participation. For content creators, journalists, civil society organisations, small businesses, and other users, access to these platforms also form part of their profession and livelihood.

Over the past few weeks, we have come across multiple counts of content restrictions on X. There have also been recurring reports from users highlighting instances of shadowbanning. This has resulted in reduced visibility of posts and accounts, unexplained reductions in content reach, and restrictions on direct messaging functionality. We have recorded a few such instances, and the details of the same can be found in the table **appended** to this letter.

Journalists and news outlets are 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, X'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.

While we acknowledge that X has, in certain instances, adopted practices that provide greater transparency to affected users regarding government requests for content restrictions, transparency regarding the existence of a request alone may not be sufficient to address the consequences of content and account restrictions. Users must also have meaningful safeguards where enforcement results in the withholding of content, restrictions on accounts, or the permanent or indefinite loss of access to an account. In particular, users should be able to understand the nature and scope of the action taken against them and have access to an effective mechanism for review.

At the same time, ordinary users may not always be aware of or able to effectively access X's reporting, appeal and grievance mechanisms. The frequent evolution of platform interfaces and user journeys, coupled with limited and sometimes inadequate information explaining how users can report an

enforcement error or seek review, can create an additional barrier to obtaining redress. Users should not have to navigate changing platform interfaces or possess a detailed understanding of X's internal reporting processes to exercise their right to challenge an enforcement action. X should therefore provide clear, accessible and easily identifiable information on how users can report, appeal and seek human review of content and account restrictions.

The consequences of erroneous or disproportionate enforcement can be particularly significant for users whose livelihoods, journalism, civic participation, or ability to communicate depend upon the platform. 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.

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 X'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 X'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. X should adopt a standardised notice that provides this information in a clear and accessible manner.

2. **Transparency regarding legal requests:** X should continue to provide transparency regarding government and court requests for content restrictions, including, to the extent legally permissible, information about the nature and legal basis of such requests, the content or accounts affected, and the action taken by X. Where X is legally permitted to do so, including in relation to actions under Section 79(3)(b) of the 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:** X should provide greater transparency regarding the role of automated systems in content moderation and account enforcement 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. X should also assess whether its automated systems adequately account for linguistic, cultural, regional, gender and other forms of diversity among users in India, particularly where context is necessary to distinguish legitimate reporting, commentary, satire or political expression from prohibited content.

4. **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 X's policies allegedly violated.

5. **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.

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

7. **Visibility, reach and distribution restrictions:** X 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 X 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.

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	Link
1.					

sflc.in

sflc.in

sflc.in