FINISHED FILE

APRIGF 2026

OCTOBER 7, 2026
APPEALING HATE SPEECH- USING THE GRIEVANCE APPELLATE COMMITTEE TO HOLD PLATFORMS ACCOUNTABLE
KANCHENJUNGA HALL
12:10 P.M. NPT

Services Provided By:
Caption First, Inc.
P.O. Box 3066
Monument, CO 80132
719-941-9557
www.captionfirst.com

This Text, Document, Or File is Based on Live Transcription. Communication Access Realtime Translation (CART), Captioning, and/or Live Transcription Are Provided in Order to Facilitate Communication Accessibility and May Not Be a Totally Verbatim Record of the Proceedings. This Text, Document, Or File is Not to Be Distributed Or Used in Any Way That May Violate Copyright Law.

>> NILAY SHAH: Hello, I'm Nilay Shah. We will talk about Appealing Hate Speech- Using the Grievance Appellate Committee to Hold Platforms Accountable. How many of have filed a grievance or reported content on a platform like Meta or anything before? Has the platform ever given a decision in your favour? Not always.

That is the decision we were facing when we were trying to appeal and report certain kinds of hate speech that shouldn't been on the platform per its own terms of service.

In India we have the Grievance Appellate Committee, it is an online form of dispute Resolution, it is a quasi, judicial body that passes binding orders on a platform.

If you see content on a platform that you don't like, you report it, the platform doesn't take it down, you have the option to appeal to the Grievance Appellate Committee, and it will pass a binding order that the platform will have to apply with.

If you are aggrieved by a decision, if it is content that you want taken down, content of yours incorrectly taken down, someone harassing you and the platform not taking the decision in your favour, you can approach the Grievance Appellate Committee.

We have seen that most content moderation on platforms is automated. That means there are no humans in the chain of grievance redress at all.

Users are circling around the same AI chat bot trying to get grievances addressed and make the case without the platform responding. That is by design. They don't want human rights issues and they want to have the money and have the LLM give you the same regurgitated responses.

The Grievance Appellate Committee provides a human in the Chan that is welcome for users.

Now that we know there is this one body called Grievance Appellate Committee in India, it is a code, a system set up under the DSA. How is a Grievance Appellate Committee filed? You as a user can appeal to the platform. If the issue is not resolved, if you have the Officer that every platform needs residence in the country. Called grievance Officer. There San online form or address where you reach out to the grievance Officer. They have to resolve the complaint in seven days.

If the grievance Officer passes an order not in your favour you can file an appeal in the GSA later. You can see with the initial levels of the grievance addressal that is reporting to a platform and to the grievance Officer, users invariably get automated responses that don't address the concern at all. That leaves no option for them but to either go to go and engage in lengthy litigation process or approach the GSA which is free of cost and wholly online. You don't need a lawyer to file your appeal either.

The context for this is that we saw that there was a lot of violent cooperate on platforms, the centre for the organized hate. The Islamophobic content on the and people remember being beat up, and donations were sought for the Groups beating up the people. This was clearly against the community standards Meta has. When we filed reports, Meta didn't take down any of the content we reported. We reached out to the grievance Officer, no dice. This was violence goal and hate content, there were 10 posts on Instagram. 50 posts during Facebook that were uploaded during an election cycle in India. All of these were explicitly denigrating a particular community. Right at the edge or over the line of what constitutes hate speech. There were subtle indicators that are content specific that were used to denigrate them and great hateful in nature. We have the speech that is proliferating and no recourse. We decided to test out the Grievance Appellate Committee. Specifically to see whether the violent and hateful cooperate could be taken down in this manner.

So the methodology that we followed was first identifying the videos, reels or posts that would actually squarely fall

under Meta's violence health speech and exploitation attacks.

We reported under multiple accounts and none were ever taken down pop directly on the platform, we tried to engage in the in-app appeal process. We didn't get a positive response. We reached out to the grievance Officer, we reported it, gave them reasoning, providing links to the offending posts and provided links why it is offensive, and give a proper legal justification of why it is violating the IT. rules and act in India and Meta's own standards. We got automated replies back. Within 30 days we filed appeals at the GSA.

In total, we filed 60 appeals specifically on this project. 10 posts on Instagram and 50 posts on Facebook that we changed. 60 to 70 posts were reported on platforms. 10 under different platforms under varied names and on telegram for example, and LinkedIn for example.

Of all the 70 posts we reported on platforms, none of the content was taken down in the first stage. All 70 posts were automatically rejected by grievance Officers. It left us no choice but to approach the GSA. Out of the five appeals that we filed at the GSA level on posts on Instagram, this was a total of 10 posts, five appeals. Seven posts were taken down. Three were not.

Out of the 50 Facebook appeals we filed.

15 posts were taken down. 10 posts were asked to be reported again on the platform itself. 25 appeals have been pending for four months at this point. The GSA researched at the outset decided it would resolve every single complaint in a span of 30 days. We have seen posts pending for 120 days and no clarity as to when they will be actioned.

When we started out this project, I think in August of 2025, we had a lot of success initially. We saw most of the posts were actioned. We had a decision quickly even in situations where we might have made a mistake or people filing appeals on behalf of us made a mistake. We got a response that said to report cooperate again, or you know, file a grievance of the said report again, for example.

Over the last couple of months, particularly since February of 2026, we have seen the dependency of the GSA has slowed significantly. Today, 10 months after, there are 50,000 appeals pending at the GSA level. Tib this is empirical research we have done. Of the appeal outcomes I have selected two that are interesting. These are Instagram appeals that we filed. Yeah.

This is one where the GSA, the point I think things started to go down at the GSA level as well. When they started to get more appeals than they could handle themselves. They started to pass orders that were unreasoned and irresponsible honestly. To file a successful appeal you need to approach a grievance

Officer and provide an acknowledgment and rejection from them. We provided both of these in the appendixes with the appeal. The platform can see the contention and see the evidence you are presenting and can submit the counter to the GSA. You as a user can never see what the platform is contending.

In this situation, Meta contended that our appeal was premature, that we had not specifically cited the sedition and didn't challenge a particular decision which is why our appeal was not maintainable in the first place.

Despite us providing them with all of the evidence they just heard Meta's contention and rejected our appeal. Next slide, please.

In another appeal, where we reported five posts, Meta very conveniently took down all five posts, two days before the GSA was to pass its order. So while we did manage to get hateful content taken down we did not get an order from the GSA mandating that Meta should take down the post. What platforms are doing instead of complying with the order, which is negative and not in their favour. In the situations where they see they're in the wrong, they will take the down. They will say we have taken this down at the grievance Officer. There is no need for an appeal or order to be passed. Next slide.

So we had three posts with Meta made contentions we could never respond to. Because the GSA can be a body that can actually be a busy influence in the country and help to take down hate speech and violent content and make the Internet a slightly more safe place for everyone. We decided to challenge this. We challenged three particular orders. Notice was issued to the Union Government and GSA. Specifically that it was the violation of the principles of natural justice. Since users never had a chance to responds to contentions that platforms made. It violated the right. And the lack of transparency around the functioning of the GSA and lack of procedural safeguards where complaints were rejected on the grounds to they were similar in nature to other appeals that artificial intelligence was very much used by the GSA to fight complaints in the first stage we decided if they're using the forms of automated grievance redressal it is only fair they publish it online and provide users the chance to see the entire procedure. That is the relief we sought.

This is still going on. Yeah, we hope to have a positive response soon.

On statistics at the GSA -- can you go one slide up, please? Yeah. I have been working with the GSA on the GSA since 2025 January.

And these stats show that there is a marked increase in the number of complaints. As of 1 January 2025, there were 230022

appeals received and 2081 dispose said. Pendency of 241 appeals in a month. That is minuscule number when you look at other courts, the quasi-judicial bodies in India inspect approximate November 2025 we had 5989 appeals. 5815 appeals disposed. Pendency of over 174. I would say until February of this year, February of 2026, we were seeing great responses.

But as of 22 September 2026, we had 63 three 50 appeals received. Only 30375 appeals disposed. Pendency of almost 50,000. The efficacy of the GSA is coming into question now. Next slide.

The time that I submitted this proposal we had seen significant successes.

It is only in the last two months we have noticed that most of the appeals we filed on hate speech are not coming through in the first place because we don't have a decision, we can't challenge the orders further in the Court of law. It leaves us and users in a state of limbo to proliferate love hate speech, it allowed it to exist permanently on the Internet pretty much.

Next slide, please.

So some of the gaps that we have identified that we would like to see the GSA remedy, the lack of transparency where users cannot see contentions the platforms raise. The GSA is going through the motions and not considering the evidence submitted and just listening to the platforms.

Delays by the grievance Officers that extend up to three months. Platform decisions being reversed just before a negative order is to be passed. Significant delays where we see appeals pending up to four months and one cannot approach court until an order is passed.

Yeah, that is the last slide. The potential fixes we have identified so far that users should be able to view and respond to platform contentiousness. There should be increased transparency at the GSA level. This is the imbalance when compared to the lawyers and unlimited resources. It shows that they don't require lawyers themselves to file appeals. Increasing the benches of the GSE, handling all of the complaints, it is impossible. If it continues, no new orders will be passed for a year. This would not only increase a precedent but increase predictability.

I will now open for questions. If any from the room and in Zoom?

>> HOST: Thank you for your detailed presentation and sharing the number and your experience while lodging complaints.

For the Civil Society, those who are working in South Asia in the field of digital democracy, what would be your top recommendation when there is no appeal or GSA mechanism and effective, I would say an effective appeal and GSA mechanism in

their particular country in South Asia and any other Region?

>> NILAY SHAH: When you don't have an appeal system you are left in an asymmetric system. You have a platform that doesn't need to listen to you and you as a user have no recourse but to engage with the platform through the human rights and safety teams to get the order passed in your favour. If you see the injustice or hashtags proliferating. That is the situation we were stuck in.

India is not the only country with out of court mechanism. The EU as out of court mechanisms. This is something that platforms can be made aware about. A situation where a binding order is passed in the platform does help create a safer Internet. I suppose in this situation, my advice would be just to try to volumetric manner reach out to the platform and get your grievances resolved. We have seen in the past, even in India, it doesn't even work out for us. Having better laws a change in the entire system of the grievance redress online would help. What would also help is lobbying platforms to have humans in the loop of grievance redressal, when it is LLM they cannot make it on your behalf. The human present will be aware of the context. Even within India you see what happens in the North of India is probably not what is happening in the South. Context and hate speech differs from Region to Region. Having humans aware of the context is helpful as well.

Anyone else? If not, thank you so much. I think I'm at time. Yeah. One minute left. Yeah. I don't think there is a question in the chat right now.

>> HOST: Thank you for the flash talk. Next up here in this room, actually right now at 12:30 is the MSG Debrief. So if you are an MSG member stay behind for the debrief. If you are not part of the MSG you are welcome to observe or head on down for lunch.

After lunch we will move to the auditorium on the ground floor for the remainder of the APrIGF sessions afternoon at 1:30 p.m., we will be the open dialogue Roundtable on the role of APrIGF in implementing the permanent mandate of the IGF. Next up is the MSG Debrief in this room and also for lunch.

(Concluded)

This Text, Document, Or File is Based on Live Transcription. Communication Access Realtime Translation (CART), Captioning, and/or Live Transcription Are Provided in Order to Facilitate Communication Accessibility and May Not Be a Totally Verbatim Record of the Proceedings. This Text, Document, Or File is Not to Be Distributed Or Used in Any Way That May Violate Copyright Law.