



## Countering hate in the digital world

Content moderation should be considered a late-stage intervention. Individuals need to be stopped early in the path to radicalisation and extremist behaviour to prevent the development of apps such as Bulli Bai.



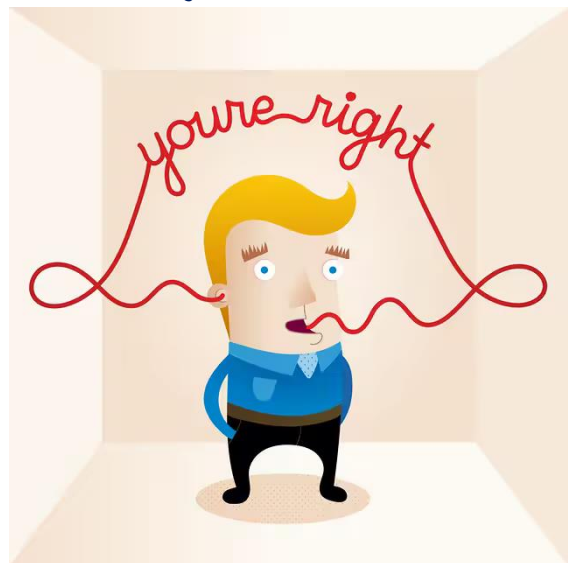
### Context:

1. Ongoing police investigations to identify the culprits behind the condemnable “Bulli Bai” and “Sulli Deals” apps, which “auctioned” several prominent and vocal Muslim women.
2. At first glance, this indicates that digital natives are not resilient against problems such as disinformation, hate speech and the potential for radicalisation that plague our informational spaces.
3. But placed within the broader context of decreasing levels of social cohesion in Indian society, such apps were even created requires us to frame our understanding in a way that can point us towards the right set of long-term interventions.



### Impact of technology on collective behaviour, and identities

1. To understand how we got here, we need to start by looking at the effect of new media technologies developed over the last 20 years on our collective behaviour, and identities.
2. Technologies have changed the scale and structure of human networks; and led to the abundance and virality of information. Social scientists



17.01.2022

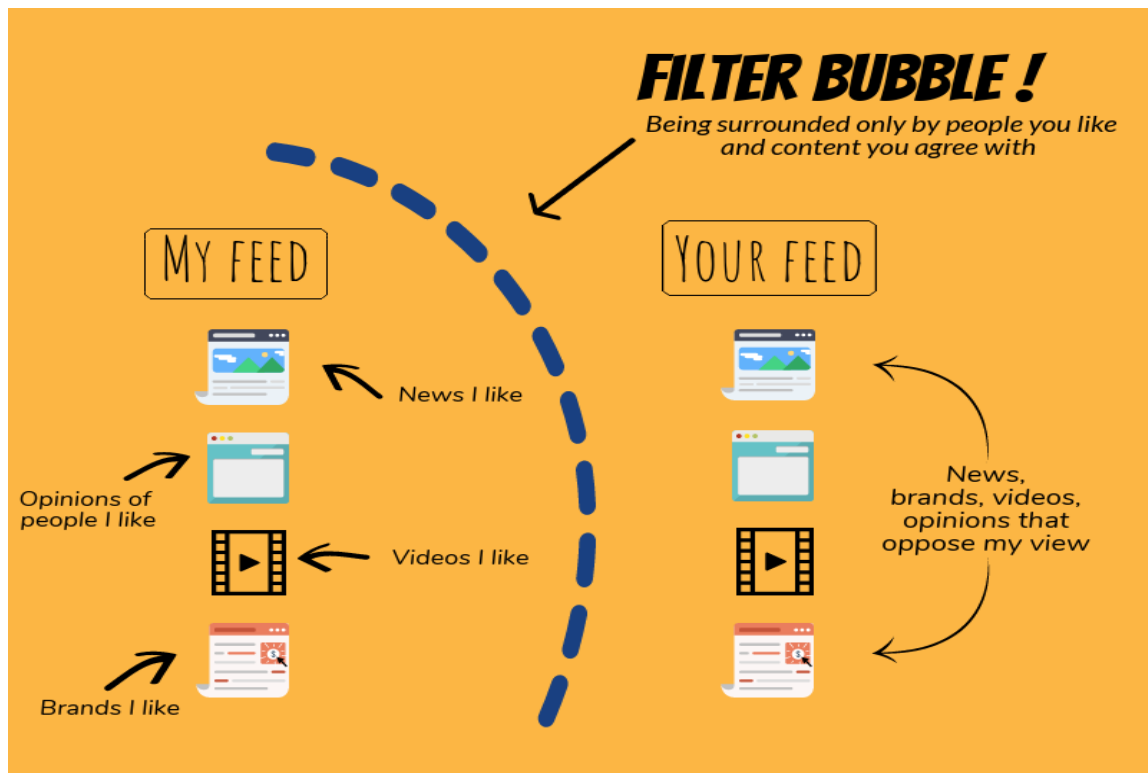
Monday



<http://www.sriramsias.com>

*hypothesise that these rapid transitions are altering how individuals and groups influence each other within our social systems.*

- 3. The pace of technological evolution coupled with the speed of diffusion of these influences has also meant that we neither fully understand the changes nor can we predict their outcomes.*
- 4. Others have focused on their effects on the evolution of individual, political, social, cultural identities. These identities can be shaped consciously or subconsciously by our interactions, and consequently, affect how we process information and respond to events in digital and physical spaces.*



Your filter bubble is your own personal, unique universe of information that you live in online. What's in your filter bubble depends on who you are, and it depends on what you do. But you don't decide what gets in - and more importantly, you don't see what gets edited out.

— Eli Pariser —

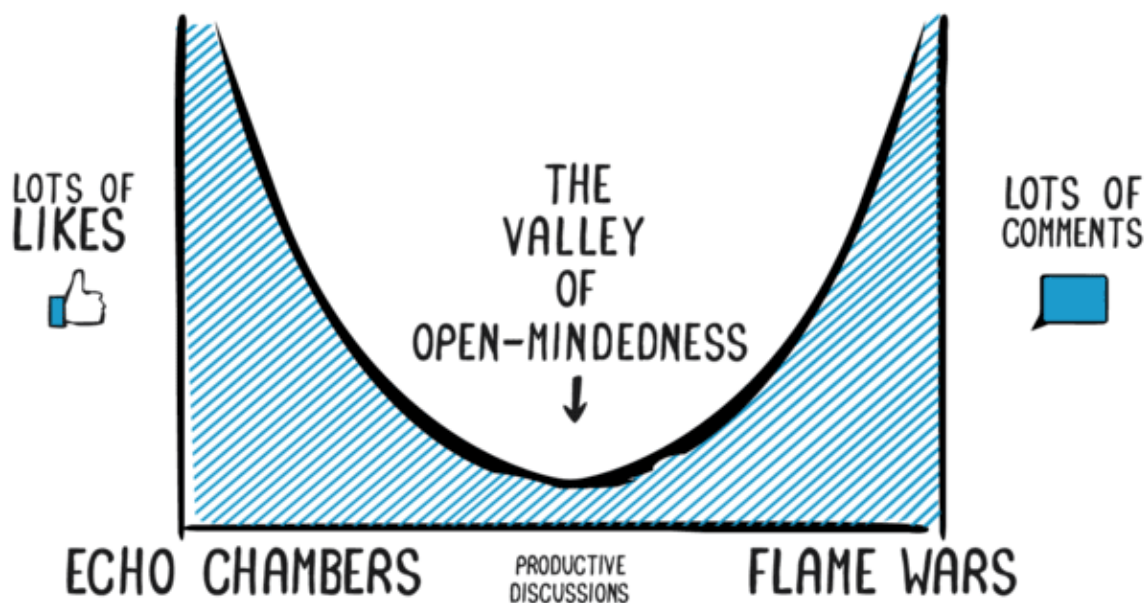
AZ QUOTES



## Digital Echo chambers:

1. Our identities ultimately bear on our cognitive processes — arguments against our defining values can activate the same neural paths as the threat of physical violence.
2. The rise of social media has been linked to the strengthening of personal social identities at the cost of increasing inter-group divisions. Some have suggested that personalised feeds in new media technologies trap us in “echo chambers”, reducing exposure to alternate views.
3. While other empirical work shows that people on social media gravitate towards like-minded people despite frequent interaction with ideas and people with whom they disagree. People can also self-select into groups that reinforce their beliefs and validate their actions.
4. Experience, though, suggests that when these beliefs are prejudices and resentment against a specific group of people, the feedback loops of social confirmation and validation can result in violence.
5. Even pockets of disconnected actions, when repeated and widespread, can destabilise delicate social-political relations built over decades.

## POLITICAL DISCUSSIONS ON THE FACEBOOK





## **Information ecosystem and pollution:**

1. Harms arising out of escalating levels of polarisation and radicalisation are primarily analysed through the lens of disinformation and hate speech which gives primacy to motives.
2. This framing leaves room for some actors to evade responsibility since motives can be deemed subjective. And for others to be unaware of the downstream consequences of their actions — often, even those taken with good intentions can have unpredictable and adverse outcomes.
3. The information ecosystem metaphor compares the current information dysfunction with environmental pollution. It encourages us to prioritise outcomes over motives, in that we should be concerned with how it spreads and not whether someone intended to pollute or not.
4. It also makes us understand that the effects of pollution compound over time, and attempts to ignore, or worse, exploit this pollution only exacerbate the problem — not just for those victimised by them, but for everyone.

## **Diverse polluters in the Information ecosystem:**

1. Our focus tends to be on those who command the largest audiences, have the loudest voices or say the most egregious things. While important, ignoring or downplaying the role of everyone else, or envisioning them as passive, malleable audiences risks overlooking the participatory nature of our current predicament.
2. Big and small polluters feed off each other's actions and content across social media, traditional media as well as physical spaces. The distinctions between "online" and "offline" effects or harms are often neither neatly categorisable nor easily distinguishable, "online" harassment is harassment.
3. Actors as varied as bored students, local political aspirants, content creators/influencers, national-level politicians, or someone trying to gain clout, etc. engage throughout the information ecosystem.
4. Their underlying motivations can range from the banal (FOMO, seeking entertainment, fame) to the sinister (organised, systematic and collaborative dissemination of propaganda, hate) to the performative (virtue signalling, projection of power, capability, expertise), and so on.
5. The interactions of these disparate sets of actors and motivations result in a complex and unpredictable system, composed of multiple intersecting self-reinforcing and self-diminishing cycles, where untested interventions can have unanticipated and unintended consequences.

17.01.2022

Monday



<http://www.sriramsias.com>

## Way Forward:

1. Several have called for action by platforms to address hate speech. Content moderation should be considered a late-stage intervention.
2. Individuals need to be stopped early in the path to radicalisation and extremist behaviour to prevent the development of apps such as Bulli Bai.
3. This is where steps such as counterspeech — tactics to counter hate speech by presenting an alternative narrative — can play a role and need to be studied further in the Indian context.
4. Counterspeech could take the form of messages aimed at building empathy by humanising those targeted; enforcing social norms around respect or openness; or de-escalating a dialogue.
5. Since behaviours in online and physical spaces are linked, in-person community action and outreach can also help.
6. Social norms can be imparted through families, friends and educational institutions. “Influencers” and those in positions of leadership can have a significant impact in shaping these norms. At such times, the signals that political leaders and state institutions send are particularly important.

## SOME KEY TAKEAWAYS AFTER REFLECTING ON TOPIC 2



### TO ESCAPE FROM ECHO CHAMBERS :

GET COMFORTABLE WITH DISAGREEMENTS AND EXPOSE YOURSELF TO CONTRASTING OPINIONS ONLINE



### OTHER WAYS TO VERIFY THE AUTHENTICITY OF ONLINE INFORMATION :

-CHECK THE DOMAIN  
-READ BEYOND SENSATIONAL HEADLINES  
-USE FACT CHECKING WEBSITES



### TO AVOID FAKE NEWS :

-BE FAMILIAR WITH THE TYPES OF FAKE INFORMATION ON INTERNET