



VIOLATION EVEN AFTER DEATH:

*THE CASE OF MUNTAHA ZAHRA AND HOW AI
TOOLS ENABLE GENDER-BASED VIOLENCE
SPILLOVERS ONLINE*

About

Digital Rights Foundation (DRF) is a women-led, not-for-profit organization based in Pakistan, working since 2013 to advance digital rights, freedoms, and online safety for all. While our work centers on the lived experiences of women and gender minorities, we actively support and collaborate with religious minorities, human rights defenders, journalists, and civil society organizations across the region. We believe in the power of a free and secure internet, one that enables expression without fear, protects personal privacy, and resists censorship. At DRF, digital rights are not abstract ideals; they are the foundation for dignity, equality, and justice in the digital age. Through direct interventions, policy advocacy, and strategic partnerships, DRF delivers digital security training, provides emergency assistance, and builds resilience among those most at risk. We work on the frontlines to combat technology-facilitated gender-based violence, promote access to information, and strengthen protections for those under threat, especially in marginalized communities. As surveillance intensifies and digital spaces become more hostile, DRF confronts both state and corporate actors through research, monitoring, and rights-based advocacy. We challenge repressive policies, expose violations, and propose practical, rights-affirming solutions. Our goal is clear: to create digital environments where safety, agency, and freedom are not privileges but guarantees for everyone.

Contact information:

info@digitalrightsfoundation.pk

www.digitalrightsfoundation.pk

Acknowledgements

This case study examines the rise in generative-AI content and TikTok accounts in the aftermath of the the rape and murder of 7 year old Muntaha Zahra in Sargodha. In the midst of an ongoing investigation and limited information, social media platforms, particularly TikTok, saw a wave of AI-generated videos that sensationalized and dramatized the incident alongside portraying narratives of misogyny, whitewashing, gendered disinformation to fuel online engagement. Through a review of the disturbing, violent and false commentary made about the incident, this report exposes how digital spaces in Pakistan are increasingly weaponized to trivialize the violence that women face and turn such incidents into an opportunity for gaining personal benefit, either in the form of sexual gratification, or financial gain.

This case study underscores the urgent need for social media platforms to:

Strengthen

Content moderation in regional languages,

Recognize

Act upon cultural nuances in online abuse, and,

Implement

Comprehensive protections for women and marginalized communities.

Authors: Maria Nazar (Research Associate), Sara Imran (Research Associate)

Design and Layout by: Ahsan Zahid and Talha Umar

Introduction

On 22 June 2026, seven year old [Muntaha Zahra](#) was found raped and murdered on the third floor of a kiryana store in Sargodha, Pakistan. According to the First Incident Report (FIR) filed by her father, Muntaha had gone to the store in the morning and never returned and upon looking for her, was found unconscious on the floor with the suspect running from the scene after throwing away a knife. In the aftermath of the investigation, two suspects were found guilty: the shopkeeper and an apprentice of his, the latter of which was killed in a 'police encounter' before he could be captured and tried in court.

This case has just been a drop in a sea of gendered violence that has gripped Pakistan over the past six months, and at the time of this report, two more cases have emerged with ever more terrifying details. Just last week, a [toddler died](#) in Karachi as a consequence of being sexually assaulted by a 17 year old boy while in a separate incident, a woman was raped in the [police station](#) by an officer who allegedly picked her on his motorcycle and brought her to the premises.

In July, two foreign women were [gang raped in Lahore](#) during a visit to Pakistan, a case that garnered national and international news coverage. The suspects, one of which was found to be a relative to a government minister, had kidnapped the women and kept them captive for several days before being apprehended by local police.

While these incidents are horrifying in and of themselves, the abuse does not stop here. With society fully functioning in two places at once, i.e. online and offline, the hatred, harm, and harassment that women face goes beyond physical violence and more often than not beyond the incident itself. These trends and patterns are only exacerbated with the introduction of artificial intelligence (AI), specifically generative AI, in the mainstream making it easier for social media users to use online platforms to further debase and abuse victims of gender-based violence.

In the fallout of Muntaha's death, the subsequent investigation, and sudden death of the alleged suspect, individuals took to social media to post AI-generated content about Muntaha, the suspect and distorted facts about the event that took place. With the use of various generative AI tools online, unprecedented levels of content was being made and posted online with the likeness of Muntaha Zahra and her suspected killer alongside elaborate, extravagant, glorifying and false narrative short-form videos.

This is not an isolated trend. Last year, after the brutal murder of [Sana Yousaf](#), a prominent internet personality, social media saw a surge in false and hateful content rooted in misogyny and which propagated existing patriarchal norms in Pakistan society. Social media users, mostly identifying as men, posted videos on TikTok, Instagram and Facebook that glorified the killer, blamed the victim, and even went as far as to generate deepfakes of Sana Yousaf “claiming” that her death was not the suspect’s fault.

Altogether identical content trends on social media were also found [later that year](#), when a young married couple from Balochistan was killed as an “act of honor.” While the incident was horrifying in itself, equally shocking was the reaction on social media where once again users took this as an opportunity to create sensational, false and deeply misogynistic posts online that dehumanised the victims. The female victim was especially targeted with sexualized content and untoward violence.

Such incidents, alongside many more, echo the reality of digital spaces in Pakistan today acting as a reflection of toxic, misogynistic and patriarchal norms of society, which often leave women and other marginalized groups exposed to harm, abuse, and violence – a trend only exacerbated by increased social media and internet use.

According to [GSMA’s Mobile Gender Gap report](#), while Pakistan has narrowed its internet usage gap since last year by 17%, millions of women are still more likely to be offline than men, usually citing affordability and safety concerns as a major barrier. Indeed, as stated by [Pakistan Telecommunication Authority’s \(PTA\)](#) last annual report, TikTok has over 65 million social media users in Pakistan out of which 71% are men. The same goes for Meta platforms, Facebook and Instagram, where more than 50% of total users are male.

Such big and nuanced digital divides become even more concerning in light of the rise in technology-facilitated gender-based violence (TFGBV) cases and the amount of gendered mis/disinformation found online. In 2025 alone, Digital Rights Foundation’s (DRF) [Digital Security Helpline](#) received 2,586 cases relating to TFGBV and noticed a rise in cases from minors between the ages of 6 to 9 – a trend that highlights severe risks to young children online today in the form of grooming, sexual abuse and exploitation.

Given these trends and the dangerous risks women and young girls face online, the response and reaction of the average social media user to the case of Muntaha Zahra comes as no surprise. DRF analysed over 60 posts and comments across TikTok, Facebook, and Instagram in regards to the murder of

Muntaha Zahra and the subsequent investigation that took place. Distortion of false, gendered disinformation and exaggerated claims about what happened and why it happened were consistently found on social media, in a larger societal narrative that glorifies killers and blames victims of abuse.



Masoom

miss.muskan19

11 Following 4976 Followers 52.6K Likes

Follow

Message

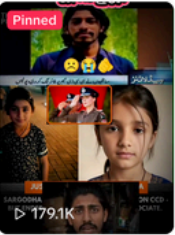
I am account creator 😊
my dream 10k followers 🥰❤️
Everyone Follow Me 🙌

Videos

Liked

Latest Popular Oldest

Pinned



179.1K

Pinned



373.9K

Pinned



582.1K



79



4557

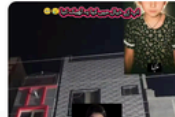
















sadiapujabpolice

sadiapujabpolice

20 Following 7240 Followers 70.4K Likes

Follow

Message

No bio yet.

Videos

Reposts

Liked

Latest Popular Oldest



2942



118.3K



4320



111.8K



3880



20.9K













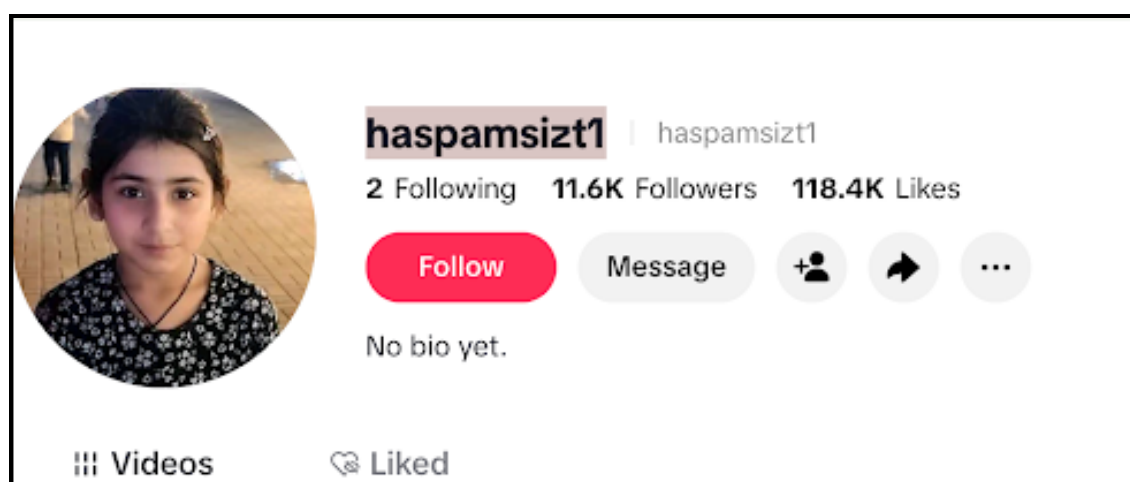
Platform specific analysis

TikTok

As limited details of Muntahra Zahra's death and its ongoing investigation were made publically available on the news, much of the general public was left afraid and confused about what took place. While most individuals took to social media to share their condolences and intense dismay at the state of gendered violence in Pakistan, others used this time as an opportunity to post various contradictory accounts online and speculate about what did or did not take place.

TikTok especially saw a disturbing onslaught of derogatory content, [with more than 20 accounts](#) dedicated to the incident, churning out at least 4-5 generative AI videos every day. These videos, created using the likeness of the 7 year old victim and the alleged suspect of her murder, portrayed sensationalised versions of Zahra's rape and murder and focused on narratives of supporting and sympathizing with the suspect in an attempt to whitewash him.

Over the course of a week, pages such as laraibqazi01, haspamsizt1, and others were cropping up that were either pivoting from their usual AI slop content to post about the incident or were completely new accounts made exclusively for the case.





media.news
media.news214

2 Following
5925 Followers
258.6K Likes

Follow
Message

💖💖💖 all about mother love 💖💖 follow and support 💖

Videos
Liked

Latest
Popular
Oldest



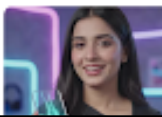














user18024335495
usereiprz88x76

6 Following
5478 Followers
38.1K Likes

Follow
Message

No bio yet

Videos
Liked

Latest
Popular
Oldest

















proud to be Pakistani
support me for 50k

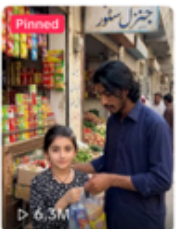
6 Following
5478 Followers
38.1K Likes

Follow
Message

No bio yet

Videos
Liked

Latest
Popular
Oldest

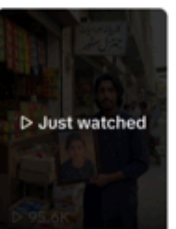








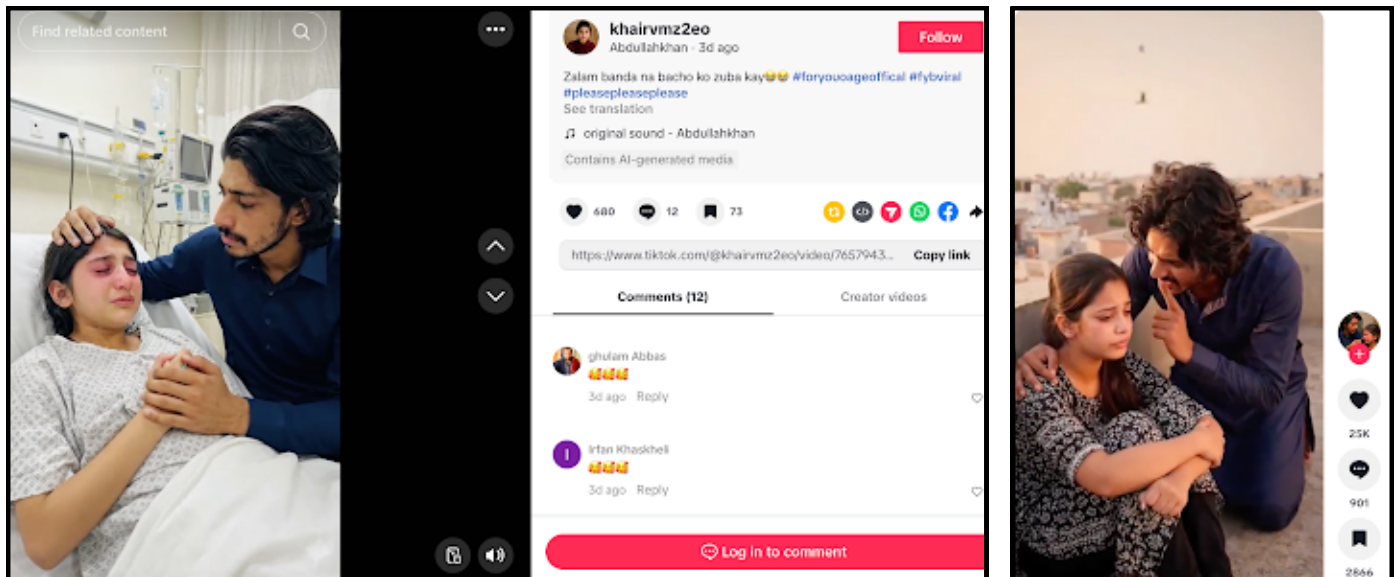



Many of the videos created and posted, showed dramatic scenes of the victim and the suspect interacting with each other, oftentimes within the context of the rape and murder that took place and sometimes in unrelated, seemingly normal interactions. Some videos, for example, showed them interacting at the kiriyana store, with the victim visiting the store to buy something and the suspect selling it to her. Others showed the victim and suspect “playing” with each other on the rooftop or in a bedroom under the guise of a sibling-adjacent relationship (with AI voiceovers where the victim called the suspect bhai or bhaiya).

Though these videos seem innocent at first glance, what they actually portray is an unsettling attempt at changing the narrative and series of events that took place. By placing these videos against the backdrop of the store where Muntaha Zahra’s murder took place, and putting normal, everyday interactions at the center of these locations belittles the crime that took place.

These videos have even more concerning and harmful iterations in the form of depictions of Muntaha Zahra as a teenager or slightly younger adult; a depiction which is far from the truth of her age and the severity of the crime. Assuming that these videos were mostly generated through freely available AI models that have limitations in quality and accuracy when it comes to facial recreations, videos that show the same content but with the victim as an older girl risks spreading disinformation that directly feeds into Pakistan’s existing victim-blaming narratives.

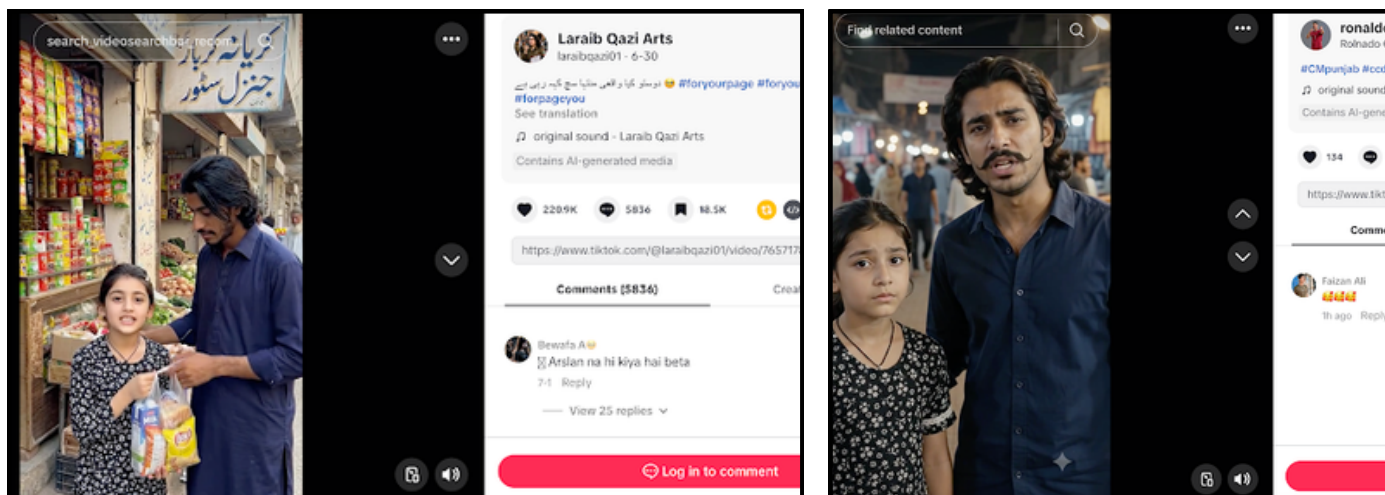


While gendered violence exerted on women and girls of all ages is equally worrying, by portraying Muntaha Zahra as a teenager, instead of as a minor 7 year old girl, gives social media users and the general public the space to impose a false sense of agency on the victim and put the onus of circumstances on her.

Another major theme that was observed throughout these TikTok accounts, involved videos where the victim and suspect were “asking the public to demand justice” for what happened to them – i.e. the murder of Muntaha and the alleged false accusation against the suspect. These videos, generated from AI tools, either showed Muntaha Zahra/the suspect addressing the screen and asking for justice or showed both of them conversing with each other, talking about how they were killed unjustly and how justice was futile because “we’re already dead.”

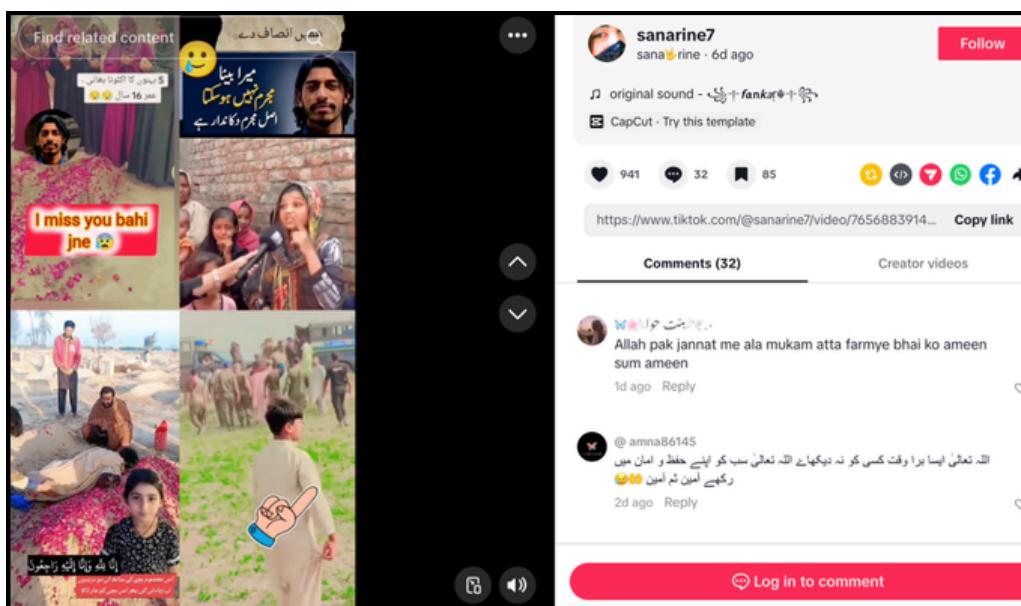
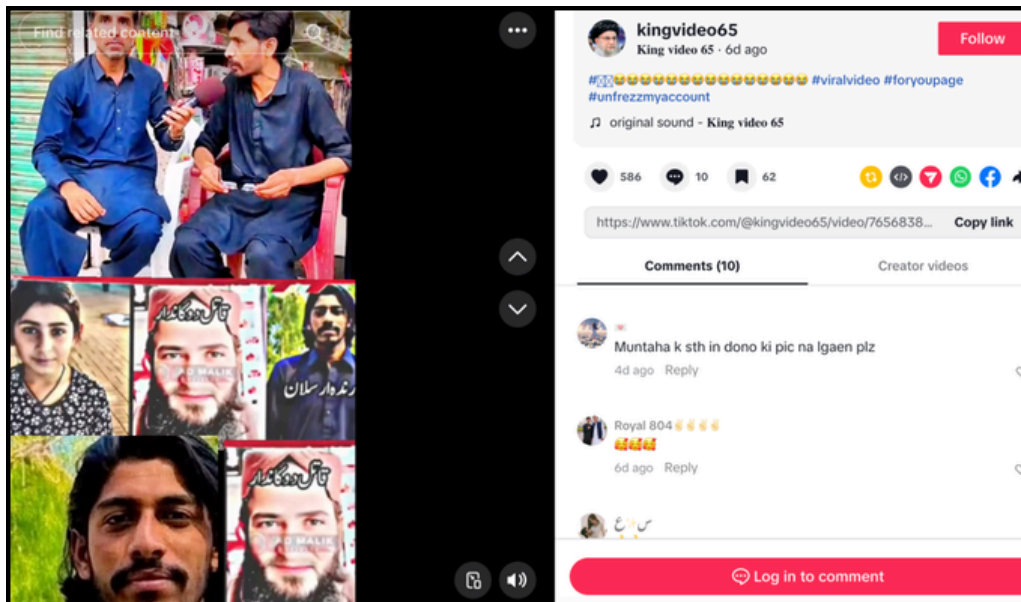
It is important to note here that these videos were being circulated at a time when unconfirmed reports were being made on social media about the investigation. According to unverified claims on TikTok accounts posing as “news accounts,” DNA tests done against Muntaha Zahra were found to be negative for the suspect – an assertion that was [categorically denied](#) by police officers and published throughout leading newspapers in Pakistan.

Multiple TikTok accounts leveraged these rumors from unreliable sources and based on little to no evidence to make dramatic, AI generated content that sympathized with the suspect, and portrayed events in a light that showed the “wrong” that was done to him. In circumstances where there is limited information available, false narratives end up doing more harm than good not only to the proceedings of the case itself, but also to the larger societal perceptions on gender violence and the state of child safety in Pakistan.

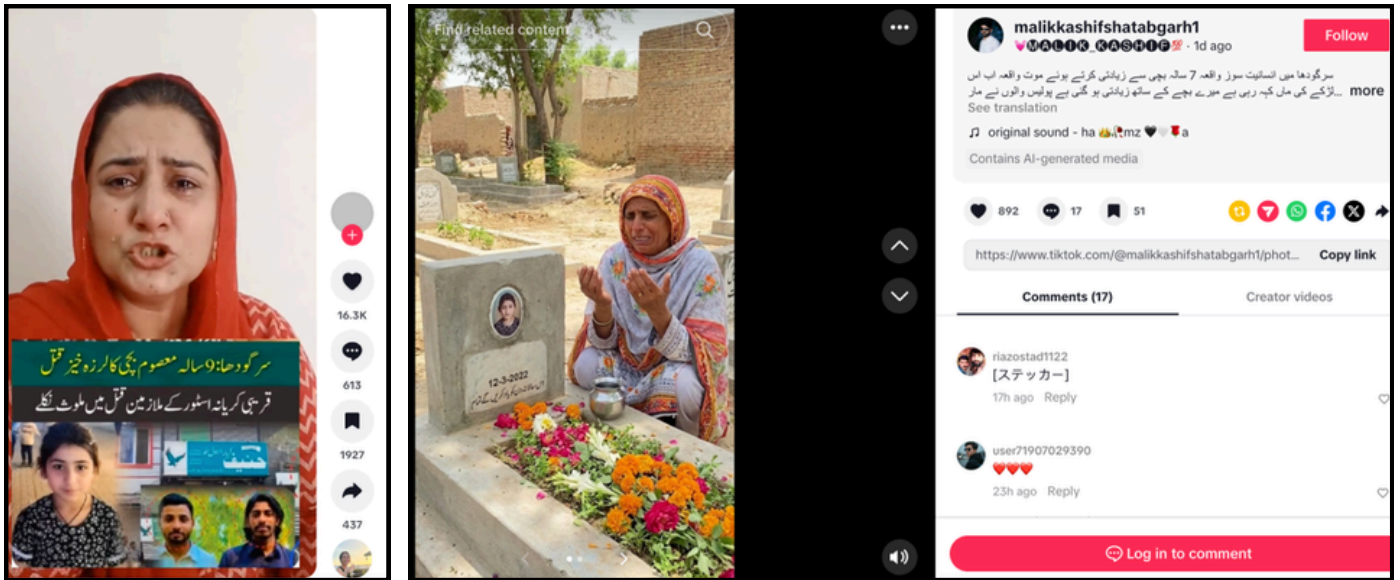


These videos, once posted, gained millions of views and likes as well as hundreds of comments where users expressed laughter and enjoyment at the scenes depicted in the videos while a few others attempted at debunking the claims being made.

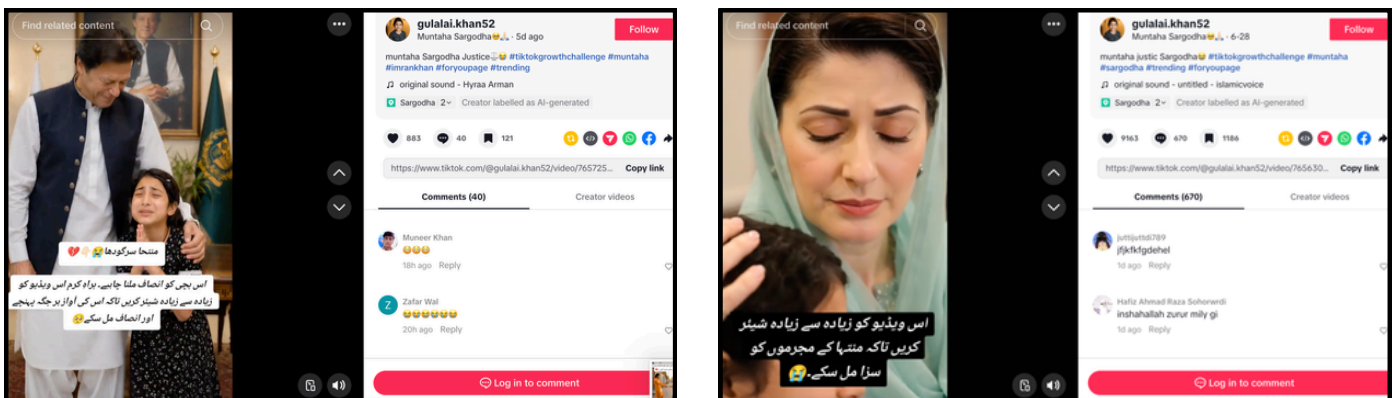
Alongside these AI-generated videos, some TikTok users who run pseudo-journalist accounts were corroborating these false claims by bringing “daily updates” to the case. Most of these videos, which featured men talking about the case and its updates with pictures of Muntaha Zahra and the suspect next to them.



In other, more worrying versions of these videos, generative AI tools were used to make fake videos of the alleged suspects' mother crying and asking for justice for her son. Sensationalizing the grief of the suspect's family was used to further narratives of how the allegations made against the suspect were false, how he was used a scapegoat for the, and how his death was a grave miscarriage of justice. In some instances, the mother was shown crying at her son's grave or at the victim's grave asking for justice.



Equally unsettling was the using this brutal murder and case of child abuse to share personal political opinions and agendas by generating AI videos of, most popularly, Imran Khan or Maryam Nawaz with Muntaha, asking for them to give the girl justice and bring her abuser into the light.



Though calling on political representatives to take action against a growing societal concern is a right of every citizen in a democratic state such as Pakistan, these videos play a part in only dramatizing the event by creating unwarranted videos and depictions of a minor girl – a violation in itself even outside of these circumstances.

Instagram and Facebook

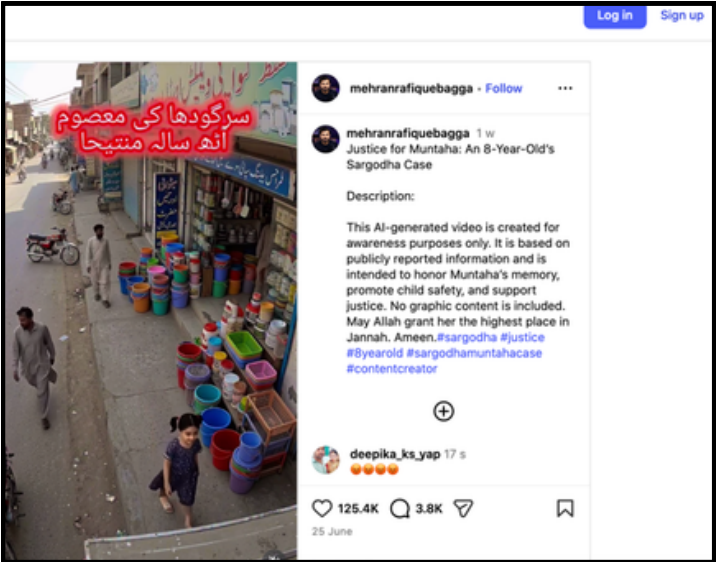
Meta platforms including Instagram and Facebook saw similar content being posted and shared but with slightly less regularity and with minor variations. While content on Meta also followed the 'sensational scripted video' route, most videos (created through generative AI models) were animated to show a little girl and grown man instead of life-like recreations of the victim and suspect.



An interesting finding made on Facebook and Instagram was the captions written under AI-generated videos of Muntaha and the suspect. Whereas TikTok videos were generally captioned with hashtags for higher reach, captions on Instagram and Facebook specifically for videos that created AI-generated scenarios of what users thought took place, provided "disclaimers" which stated that these videos were made for "awareness purposes."

One such video was particularly made from the point of view of a CCTV camera showing Muntaha walking to the kiriyana store and interacting with the shop keeper – a piece of evidence that was reported on in multiple news reports as being the only footage available that shows the victim going to the shop and never coming back.

Similar to TikTok, social media users who presented themselves as self-proclaimed journalists alongside unverified “news channel” pages were seen not only spreading disinformation through false claims but also by using dramatized versions of the incident to draw in audiences. In many such videos, men presented misogynistic opinions and victim-blaming narratives as ‘facts’ of the case.



Discussion

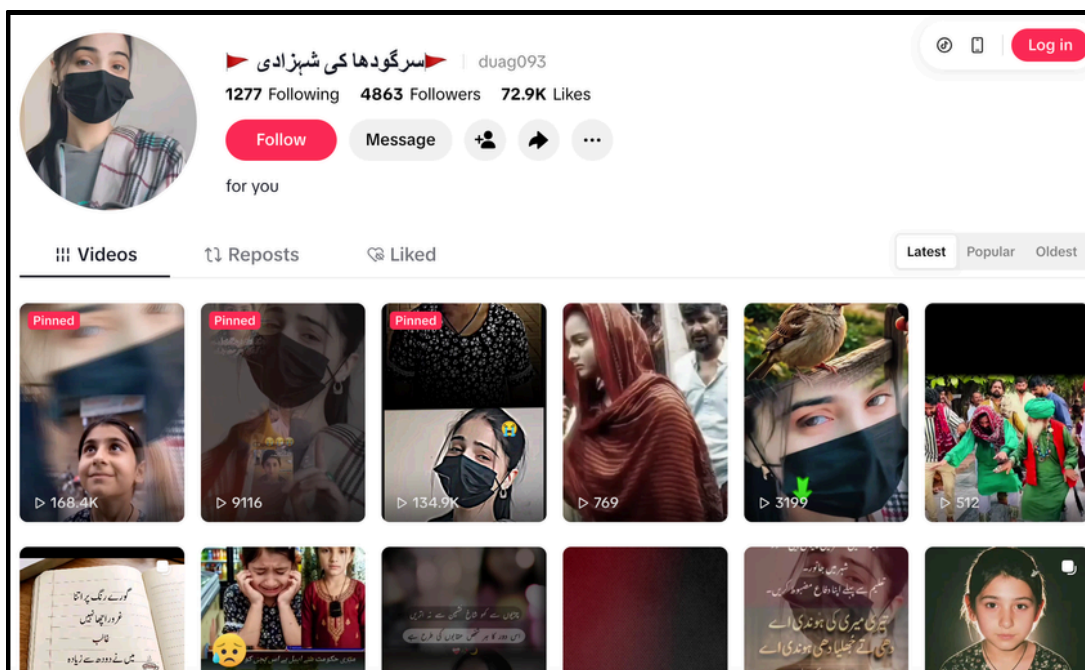
With every new case of gender violence in Pakistan that is highly publicised and covered by the media, a disturbing and recurring pattern of social media activity seems to be emerging: AI-generated content that dramatizes and sensationalizes real world violence against women and that portrays the sexualization of women. Such content has been found to either be created by inputting imaginary descriptions of what the women should look like or more often than not, are based off of real women and children found online or on the news. According to a report by [UN Women](#), “deepfake pornography makes up 98% of all deepfake videos online, and 99% of the individuals targeted are women,” out of which AI-generated child sexual abuse material takes up a chunk that is on a rapid rise.

[Research](#) shows that traditional forms of AI-generated image based sexual abuse are generally created to seek social reward, get amusement and/or recognition from their peers and commonly follows patterns of exploitation that affect women in public spaces. Indeed, as we’ve seen from in the case of Muntaha Zahra, the case of Sana Yousaf, and the cases of countless other women and girls who have been abused, harassed or exploited in the public eye, content creators capitalize on actual events, be that incidents of gendered violence, public events that feature prominent female public figures or even natural disasters, and treat horrific crimes and events as plot points and scripts, to farm for maximum engagement from social media platforms that in turn reward high yielding content and prioritize likes, comments, shares, numbers and figures, over rules and regulations of platform use and minimizing harm.

Similarly, many perpetrators of AI-generated image based abuse also cite [curiosity mixed with creativity](#) as a reason for partaking in creating AI-generated content – claiming that their behaviour was “[just a joke](#)” with them having no idea about the harm their actions could and did cause. One can only assume that such behaviour and thinking carries forward to Pakistan and South Asian societies, where men have historically used women and girl’s bodies to gain control, power and stir narratives in their favour – whether for themselves on an individual basis or for all men in general. Treating the AI-generated images of women and minors in sexualised ways as “fun” deeply downplays the harm this content causes not only to women online but especially to women offline. Social media platforms are equally responsible for pushing forward and intensifying this by rewarding such material and content. Despite AI-generated content violating the existing terms of service of TikTok, for example, accounts

escape moderation for months because of the [sheer volume of content](#) that they post – usually going up to more than 70 videos per day. [Reports](#) on Meta and X, made by whistleblowers from the company, show that harmful content (which includes content on misogyny or violence,) is actively allowed to circulate online in an attempt to ‘compete with TikTok.’ By developing algorithms that maximize profits gained from making and posting harmful, sensationalized and sexualized content, Big Tech companies are directly responsible for giving social users a financial incentive to make and disseminate AI-generated images and videos of women and young girls in misogynistic and derogatory ways.

But the blame not only lies with users who create such content and push it online, but with individual users who disseminate, share and amplify such content. Consumption of AI-generated content that shows sexualized images of women and girls or portrays gender-based violence [fulfills](#) sexual gratification similar to the kind that men receive from watching or viewing pornography. Social media users have also gone as far as to claim that viewing such images and videos on social media arouses them.



Role of platforms

Why context-specific moderation and curbing generative AI matter

Generative AI has fast become a mass driver of disinformation and exploitative content, particularly in under-resourced information environments like Pakistan, where local-language detection capacity lags far behind the volume and sophistication of synthetic content being produced. Platforms' one-size-fits-all, primarily English-language moderation systems consistently miss context that would be obvious to a local reviewer, allowing dramatised, exploitative, or misleading AI content involving real cases, like Muntaha Zahra's, to circulate unchecked while genuine reporting and survivor voices are drowned out. As the Oversight Board itself has [noted](#), deceptive AI-generated content is now a defining feature of how tragedy and conflict are represented online, and platforms' approach to surfacing and labelling this content "must evolve" to keep pace.

Where platforms are failing

Beyond isolated failures to act, platform monetisation and policy structures point to the fact that these shortcomings are inherent to the system. Investigations (including DRF's own EOBI [disinformation on YouTube](#)) have documented creators, including [from Pakistan](#), earning ad revenue directly from platforms by mass-producing divisive or exploitative AI-generated content, showing that current monetisation systems actively reward volume and virality over accuracy or dignity.

Meanwhile, the Oversight Board has repeatedly [described](#) Meta's manipulated media rules as "incoherent", covering only narrow categories such as AI-altered video of someone saying something they didn't say, while leaving AI-generated images, audio, and dramatised "recreations" largely unaddressed. Thus platforms continue to collate AI-generated content with general misinformation or manipulated-media policies not designed for it, rather than adopting the standalone framework with provenance standards (e.g., C2PA) and mandatory "high-risk" labelling that the Oversight Board has explicitly [called](#) for.

Recommendations

Platforms must actively take actionable steps to curb the spread of TFGBV and especially AI-generated content dramatising and sensationalising real crimes involving children and women in Pakistan and the Global South. Some recommendations for platforms, guided by human rights compliant policies such as the United Nations Rights of the Child (UNCRC), United Nations Guiding Principles on Business and Human Rights (UNGPs), as well as decisions made by Meta's independent Oversight Board, are provided below.

Cross-Platform Recommendations

- **Mandatory synthetic-content labelling:** Apply clear, prominent “AI-generated/dramatised recreation” labels to any content depicting real crimes or victims, consistent with the Oversight Board's call for a [dedicated AI-content policy with provenance standards](#) (C2PA) distinct from general misinformation rules.
- **Extend protections to private individuals:** The Oversight Board's finding that [non-consensual AI depiction protections must cover non-public figures](#), should be explicitly extended to private individuals and their families, including deceased minors, who cannot consent to being depicted.
- **Demonetise exploitative recreations:** Cut ad-revenue sharing for channels/accounts that build monetised content around dramatised real-crime narratives involving children, closing the financial incentive for sensationalism.
- **Family reporting channel:** Provide victims' families a fast, accessible route to flag and request removal of synthetic content depicting their case, grounded in [UNCRC](#) Article 16's protection of a child's privacy and honour.
- **Civil Society partnership:** Work with local child-rights and digital rights organisations such as [Sahil, Zindagi Trust](#), and DRF, to distinguish legitimate public-interest reporting from exploitative sensationalism in the Pakistani context.
- **Algorithmic de-amplification:** Ensure recommendation systems do not prioritise synthetic dramatisations over verified reporting during active or sensitive cases.

Platform-Specific Recommendations

TikTok

This platform has the highest-risk format given its short, algorithmically-amplified narrated videos.

- > Enforce AI-content labels specifically for true-crime-style content.
- > Trigger a rapid-response review when a real, active case goes viral to prevent algorithmic amplification of unlabelled synthetic recreations.

Instagram

- > Apply the same labelling standard to Reels.
- > Prioritise review of reported content depicting real, identifiable minors, regardless of virality.

Facebook

- > Close monetisation loopholes allowing Pages/Groups to profit from dramatised crime content about real child victims.
- > Escalate viral posts for enhanced human review.

X (Twitter)

- > Require AI-provenance labels at scale.
- > Integrate Community Notes context for unlabelled synthetic recreations.
- > Enforce against monetised accounts spreading dramatised content as if authentic.



DigitalRightsFoundation
"KNOW YOUR RIGHTS"



@DigitalRightsFoundation



@digitalrightsfoundation



@digitalrightsfoundation



@digitalrightsfoundation



Digital Rights Foundation



@digitalrightspk.bsky.social



@DigitalRightsPK



@DigitalRightsPK

digitalrightsfoundation.pk