iVERIFY PAKISTAN
AI, MISINFORMATION AND
FACT-CHECKING REPORT

2024-25

IBA CEJ
Center for Excellence in Journalism

iVerify
PAKISTAN

UNDP

iVERIFY PAKISTAN AI, MISINFORMATION AND FACT-CHECKING REPORT 2024-25

iVerify is the United Nations Development Programme's non-partisan fact-checking initiative which is being implemented in Pakistan by the Centre of Excellence in Journalism (CEJ) of the Institute of Business Administration (IBA) in Karachi. It was launched officially in January 2024 with **the mission to empower people with verified and reliable information.** The project has been launched[1] in different countries by UNDP with national counterparts to support actors around the world in the prevention and mitigation of disinformation, misinformation and hate speech.

Authors:
Zahrah Mazhar with Syed Talal Ahsan, Ainum Baig and Muzhira Amin

Contributors:
Hina Hussain, Razi Wani, Bazigah Murad, Aleezeh Fatima, Mahnoor Tariq

Acknowledgments:
Muneeba Iqbal and Masood

Designers:
Roshail Khan and Muhammad Wasi

[1] https://www.undp.org/digital/iverify

CONTENTS

1

2

3

4

5

INTRODUCTION

iVerify Pakistan was launched at a critical juncture as Pakistan prepared to vote in one of its most tumultuous general elections after multiple delays[2] and amid highly visible political polarisation among the public. Two years later, the team's first public report comes at another critical juncture: the Artificial Intelligence reckoning.

In 2026, what will the advancement of generative AI mean for the news media whose space has already been taken up by social media? What will it mean for false information given that a misleading but realistic visual or audio can be achieved with a single prompt? Can the media use AI to its advantage to reclaim some of its lost audience and will fact-checkers make a friend or foe of AI in their quest to report evidence-based truth?

This report won't answer these questions.

Most experts agree that AI's impact is anything but straightforward[3]; we need to **continuously monitor it, engage with it and try to stay a step ahead** so that the alarms that rang loudly – and are still ringing – about AI posing all kinds of risks, from taking over human jobs to killing creative and critical thinking, don't impede our ability to have technology serve our goals.

This report does just that: it presents iVerify's findings, data and analysis after monitoring and engaging with AI through its work. We didn't set out to write this report focusing on AI but as the world shifted its focus, so did iVerify's work in fact-checking. While we identified patterns, perpetrators, and types of misinformation and disinformation during 2024, AI creeped into each aspect in 2025, and set the tone for 2026.

2 https://www.aljazeera.com/news/2023/9/22/political-parties-react-to-pakistans-general-elections-announcement.
3 https://www.weforum.org/stories/2025/12/ai-paradoxes-in-2026/

Given the extensiveness of the topic of mis- and disinformation, this report can only convey a small part of the data and analysis. iVerify is a small team but we have been able to gain the interest of government bodies, such as the Pakistan Telecommunication Authority (PTA) and Election Commission of Pakistan (ECP), maintain and form new partnerships with leading media houses in Pakistan, including *DAWN* and *Express Media*, earn the trust to be quoted by publications with a wide reach such as *BBC Urdu*[4], *PTV* and *AFP*[5], and instil fact-checking knowledge in the next generation through its programme, iVerify Advocates.

We hope through this report, despite the project's challenges and limitations, iVerify can continue to garner support and make an impact, and assist all those navigating the AI reckoning.

---

[4] https://www.facebook.com/reel/1204097075003090
[5] https://factcheck.afp.com/doc.afp.com.36Q88EE

İVERIFY PAKISTAN
THE TEAM

PROJECT LEAD

Zahrah Mazhar is an editor and journalist with more than 15 years of experience in print and digital media. Her areas of interest include social media and news/information consumption, the impact of AI on journalism and its editorial adaptation, and helping newsrooms implement practical solutions to countering misleading and false information.

PROJECT SUPERVISOR

Syed Talal Ahsan is a graduate in Economics and Politics from the Lahore University of Management Sciences, who is currently working as a News Editor at Express Tribune. In his role as Project Supervisor, he leads the team's daily operations and liaises closely with the Project Lead to achieve programme objectives.

SENIOR FACT-CHECKER

Ainum Baig is a graduate with an MA in Mass Communication from the University of Karachi. Ainum started working with iVerify as part of the launch team and is one of the key persons who identifies patterns and trends of false information and assists in training new fact-checkers.

FACT-CHECKER

Muzhira Amin is a multimedia journalist with six years of experience covering climate change, human rights, and animal welfare. She is currently with DAWN, where she specialises in long-form, in-depth reporting, combining fieldwork, data and visual storytelling to examine structural and social injustices.

FACT CHECKER

Muneeba Iqbal is an MS Journalism student at IBA-CEJ and holds a BS English degree. She interned at Soch during the 2024 general elections, verifying election-related content. Her work reflects an interest in the electoral process, governance and public discourse.

FACT CHECKER

Masood is an MS Development Studies student at IBA Karachi. He has previously worked as an iVerify Advocate intern. He has also worked with Habib University as a research intern and with the Islamic Chamber of Commerce and Development (OIC) in its International Relations department.

SOCIAL MEDIA MANAGER

Roshail Khan is a content producer and digital storyteller with over five years of experience. With an MA in Mass Communication from the University of Karachi, he holds a strong understanding of how stories travel, shape opinion and influence public discourse in digital spaces. At iVerify Pakistan, he focuses on transforming daily fact-checks and verification processes into clear and interactive narratives while driving reach and engagement across social media.

CREATIVE MEDIA PRODUCER

Muhammad Wasi has a bachelors in Media Studies from Iqra University and has built a diverse portfolio with SOC Films and DAWN Media Group. Specialising in editorial direction and digital storytelling, he currently leads the multimedia execution and visual identity of iVerify Pakistan.
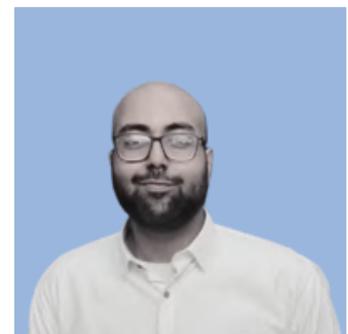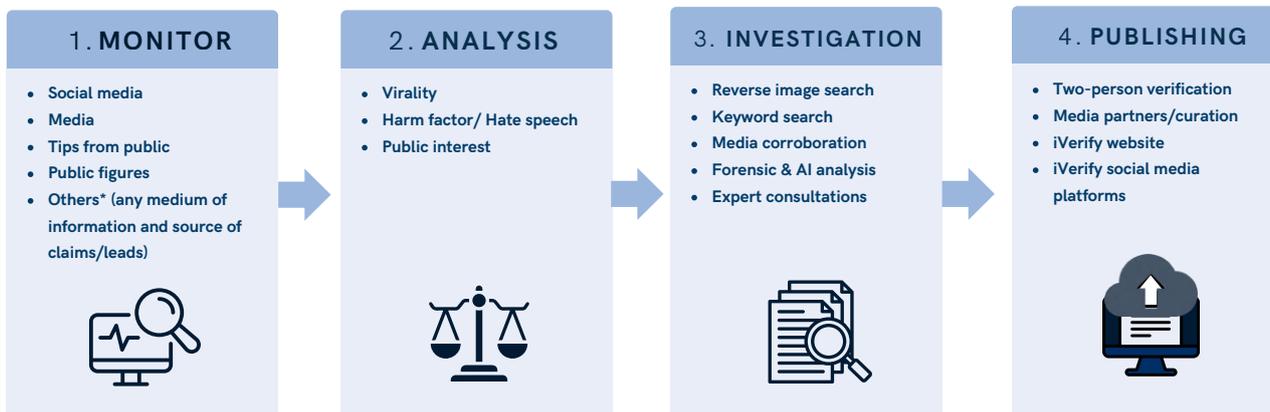
PROJECT COORDINATOR

Saroj Rizwan Khan holds undergraduate and graduate degrees in International Relations and is currently pursuing an M.Phil. Her expertise spans international law, peace studies and research, complemented by her work as a certified trainer and former academic and programme lead. She acts as the main liaison between different stakeholders of the iVerify Pakistan project.

METHODOLOGY

iVerify Pakistan follows a four-step methodology for fact-checking:

1. MONITOR	2. ANALYSIS	3. INVESTIGATION	4. PUBLISHING
• Social media • Media • Tips from public • Public figures • Others* (any medium of information and source of claims/leads)	• Virality • Harm factor/ Hate speech • Public interest	• Reverse image search • Keyword search • Media corroboration • Forensic & AI analysis • Expert consultations	• Two-person verification • Media partners/curation • iVerify website • iVerify social media platforms

How do we select claims: The claims/leads found in monitoring are evaluated for either further investigation or are discarded. The claim is reviewed to make sure that it is not an opinion.

Investigation: iVerify uses primary and secondary evidence in almost all fact-checks, comprising multiple sources, including the quoted individual, public data, thematic experts, relevant authorities and technical tools (for visuals, audio, geography, archival, etc) to ensure all sides of the story are investigated.

Criteria for claims pursued: Virality, public interest, potential harm and relevance to the ongoing news cycle. Some claims, depending on their significance in the context of news, public interest and projection of their circulation based on similar past claims, might be selected for pre-emptive debunking, despite low virality, to mitigate and control the potential damage they might cause.

Criteria for claims discarded: Lack of access to a place, documents or response from authorities; limitation in investigative resources; inconclusive findings to avoid adding to the confusion; evaluation of 'more harm vs good' by sharing a piece of unverified information – even in the form of a fact-check – when the claim can pose danger to someone's life, expose personal details, etc.

What is the two-step verification? The selected leads, based on their nature, are investigated using visual verification and forensic tools, corroborated with official sources and consulted with reporters and experts. The two-step verification means that after the fact-checker assigned to the claim has completed their investigation, another fact-checker as well as the Project Supervisor/Project Lead reviews the process to ensure accuracy before publication.

TRUE

The claim is rigorous and the content is demonstrably true.

MISLEADING

The statement contains correct data, but ignores very important elements or is mixed with incorrect data giving a different, inaccurate or false impression.
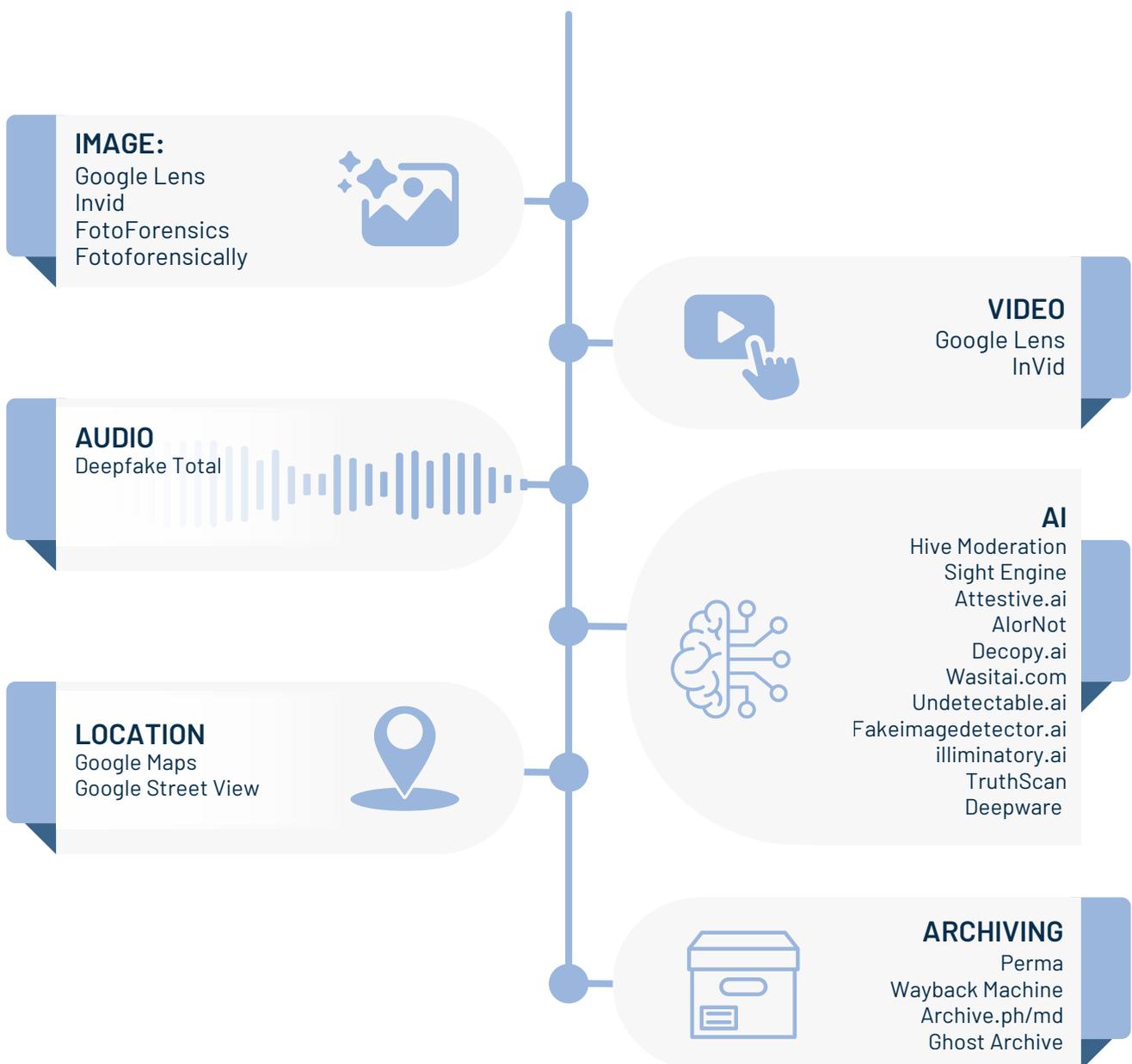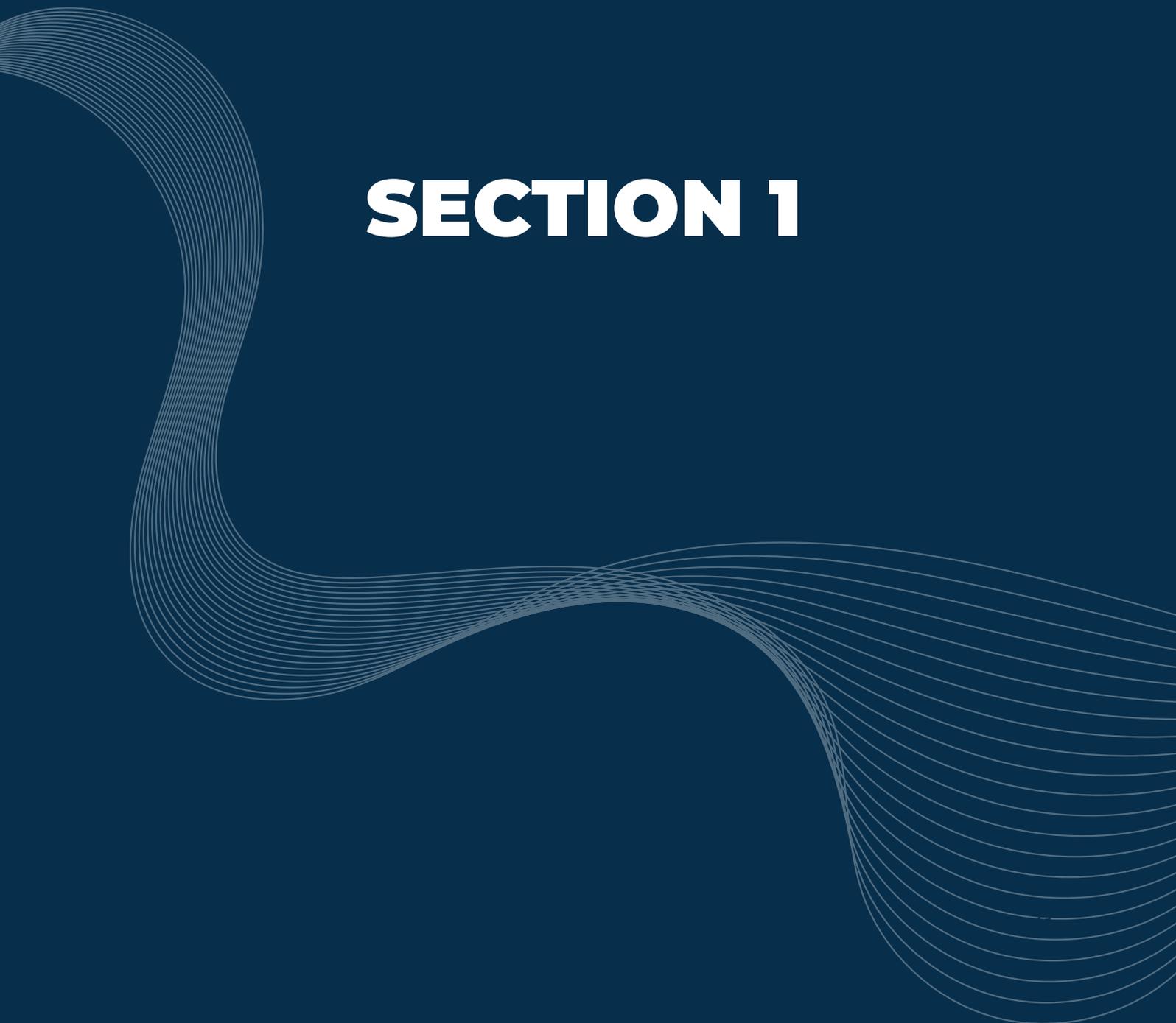
FALSE

The claim is inaccurate according to the best evidence publicly available at this time.

TOOLKIT

Tools play a key role in iVerify's fact-checking process but bearing in mind that they can also give false/inconclusive results, the team keeps using different softwares and relies on more than one to provide as evidence. Particularly for AI-generated content, tools are used as supportive aids that give likely indications about the status of a piece of content, rather than any conclusive results, and have to be used in conjunction with traditional visual and logical analysis.

IMAGE:
Google Lens
Invid
FotoForensics
Fotoforensically

VIDEO
Google Lens
InVid

AUDIO
Deepfake Total

AI
Hive Moderation
Sight Engine
Attestive.ai
AIorNot
Decopy.ai
Wasitai.com
Undetectable.ai
Fakeimagedetector.ai
illiminatory.ai
TruthScan
Deepware

LOCATION
Google Maps
Google Street View

ARCHIVING
Perma
Wayback Machine
Archive.ph/md
Ghost Archive

Some of the common tools that iVerify Pakistan uses for its investigations and digital forensics.

SECTION 1

1.1 EXECUTIVE SUMMARY

Checking facts is the backbone of journalism; fact-checking, however, is often overlooked despite social media and AI rapidly reshaping information environments across the world. iVerify was launched in Pakistan ahead of the 2024 general elections as a response to the widening gap between verified and unverified information in the form[6] of:

Misinformation: Information that is false, but not created with the intention of causing harm

Disinformation: Information that is false and deliberately created to harm a person, social group, organisation or country

Malinformation: Information that is based on real facts, but manipulated to inflict harm on a person, organisation or country

Over the nearly two-year period covered in this report (December 2023 to November 2025), iVerify monitored 1,026 potentially false or misleading claims, pursuing 513 for in-depth fact-checking across politics, religion, sports, conflicts, and social issues. This report documents the evolution, scale and impact of misinformation, disinformation and malinformation in Pakistan based on iVerify's work.

SCALE AND NATURE OF MISINFORMATION

The findings show that false and misleading content overwhelmingly dominated the information ecosystem while politics emerged as the most exploited domain, accounting for the vast majority of claims, particularly during elections, political unrest, judicial developments and periods of conflict. Social media — especially X (formerly Twitter) — served as the primary vector for the spread of mis- and disinformation while encrypted platforms such as WhatsApp remained largely unobservable despite their known influence.

[6] https://www.undp.org/eurasia/dis/misinformation

Patterns identified across the dataset reveal consistent and deliberate tactics, including:

- Half-truths that blend factual elements with falsehoods to build credibility,

- Missing or distorted context, often recycled during breaking news cycles,

- Mistranslation and linguistic manipulation, especially from foreign-language sources,

- Impersonation of journalists, media outlets, and public figures, and

- Doctored, synthetic, or AI-generated visuals, increasingly difficult to distinguish from authentic material.

These tactics demonstrate how intentional disinformation frequently mutates into widespread misinformation once amplified by broader audiences.

POLITICAL ACTORS, TARGETING, AND COORDINATED CAMPAIGNS

Analysis of active disseminators indicates that political supporters and organised networks played a central role in amplifying false narratives. Claims originating from or amplified by supporters of major political parties, alongside coordinated foreign-origin accounts, significantly shaped online discourse. These campaigns often aimed to discredit opponents, undermine institutions, suppress dissenting movements or inflame ethnic, religious and regional divisions.

The report documents how misinformation was strategically deployed to:

- Undermine electoral processes and voter confidence,

- Discredit rights movements, particularly related to Balochistan,

- Incite ethnic groups during periods of political unrest, and

- Legitimise state or societal hostility by labelling groups as "anti-state".

AI AS A FORCE MULTIPLIER

A central finding of the report is the rapid acceleration of AI-enabled misinformation in 2025. **AI-related fact-checks increased sixfold compared to 2024**, reflecting wider access to generative tools capable of producing realistic images, videos, voice clones and altered visuals.

AI was used to:

- Create deepfake endorsements, fabricated speeches and manipulated videos of political leaders,

- Spread false narratives during regional conflicts, including Pakistan–India and Iran–Israel tensions,

- Target women, particularly politicians and journalists, through sexualised and reputationally damaging synthetic content.

- This "AI-fication of reality", where authentic content is subtly altered rather than wholly fabricated, has made verification more complex and time-intensive, while simultaneously eroding public confidence in all visual evidence.

HATE SPEECH

Beyond factual inaccuracy, iVerify identified hate speech as a persistent and dangerous feature of the claims found; of 513 published reports, 34 addressed cases of explicit hate speech targeting individuals or groups based on religion, ethnicity, gender or political affiliation.

Religious and sectarian misinformation proved particularly volatile, with recycled images, misattributed foreign visuals, and emotionally charged falsehoods strategically circulated during moments of unrest to inflame tensions between communities. These narratives exploited historical grievances, deepened mistrust and posed tangible risks to social cohesion and public safety.

CHALLENGES TO FACT-CHECKING

The report outlines structural and operational challenges facing fact-checking in Pakistan, including:

- The speed and scale of viral misinformation versus slower verification processes,

- Persistent recirculation of debunked content,

- Politically insulated echo chambers,

- Internet and social media clampdowns as witnessed in 2024 during and after elections,

- Increasing difficulty in detecting high-quality AI-generated media,

- Limited access to reliable AI-detection tools,

- Overreliance on unverified AI chatbots for "instant verification" and

- Restricted access to official information.

LOOKING AHEAD

As AI-generated content becomes increasingly indistinguishable from authentic media, the report underscores that fact-checking must evolve from a reactive exercise to a sustained public-interest function. Strengthening information resilience will require coordinated action among journalists, fact-checkers, media organisations, academic institutions, civil society, technology platforms, policymakers and the public.

The findings suggest that while public scepticism toward online content is slowly increasing, trust deficits remain acute. In this environment, credible fact-checking initiatives such as iVerify are essential to preserving democratic discourse, protecting vulnerable groups and safeguarding the integrity of Pakistan's information ecosystem as the country moves into an even more AI-saturated future.

KEY RECOMMENDATIONS FOR STAKEHOLDERS

- Fact-checking must be institutionalised as a long-term public-interest function, supported by sustained funding, media partnerships and preparedness for internet and platform disruptions.

- Government bodies and public institutions need to play a central role in countering political and electoral misinformation through timely communication, modern digital outreach and coordination with media and fact-checkers.

- Newsrooms need to activate specialised response mechanisms and increased AI literacy across the board to tackle AI-enabled misinformation, including gendered deepfakes and manipulated visuals.

- Hate speech and harmful misinformation should be addressed as threats to social cohesion, demanding collaboration between regulators, media, fact-checkers and civil society.

- Social media platforms must independently strengthen accountability, particularly around political advertising, and amplification of verified information, with digital rights activists and international development bodies putting pressure on them to do so.

- An informed public is the final line of defence, underscoring the need for sustained media literacy and civic awareness initiatives.

In 2026, iVerify aims to play a pivotal role in putting higher value on accurate content and mitigating the risks of AI's sophisticated use and high social media penetration by forming strategic partnerships with

- Journalism fraternity
- Academic institutions
- Media development groups
- Government bodies

Fact-checkers, assisted by the combination of AI detection tools, editorial skills and media literacy initiatives, will be essential in building resilient information ecosystems that can preemptively counter misinformation.

1.2 KEY DATA

In this section, the iVerify Pakistan team has provided a visual presentation of the data collected during the fact-checking process, with a focus on the number of claims debunked, their nature and the perpetrators behind them.

The team observed that misinformation and disinformation proliferated across various platforms, from social media to traditional news outlets, with the former being a more active medium than the latter and others.

The data presented in this section is from December 2023 to November 2025. Certain metrics overlap and have been accounted for individually in each section in the calculation.

1. TOTAL CLAIMS FOUND, PURSUED, DISCARDED

iVerify identified a total of 1,026 leads in its two-year operations, with an even split between 513 pursued (50%) for further investigation, while 513 were discarded (50%).

**Total leads: 1,026**

- Pursued
- Discarded

50%   50%
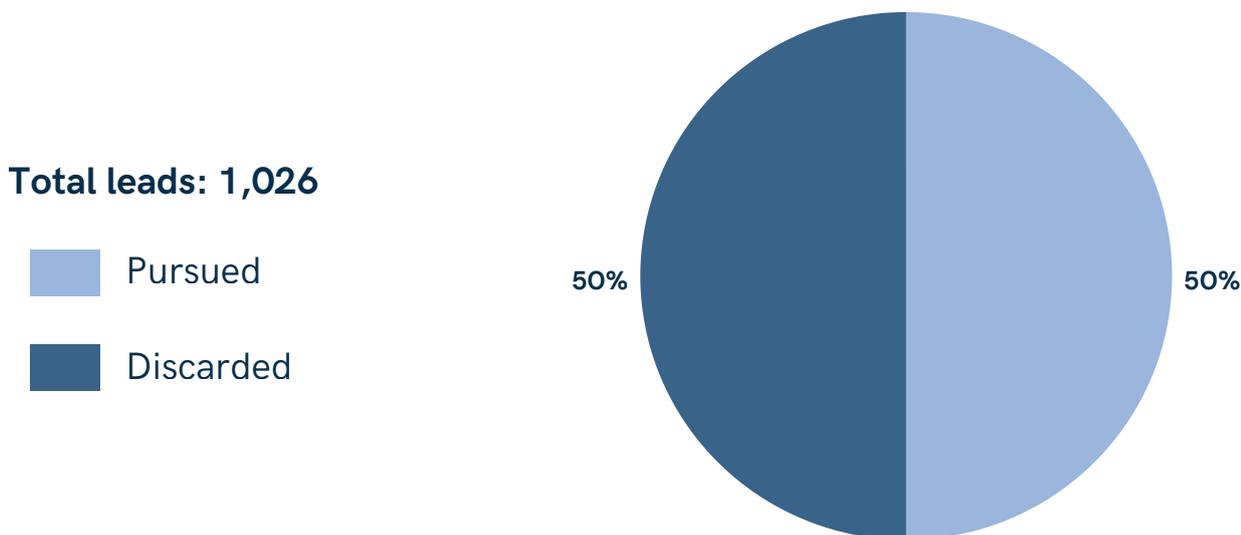
2. PURSUED CLAIMS: RESULTS

Between December 2023 and November 2025, the majority of the fact-checks carried out were found to be false (401 or 78.2%), followed by those found to be misleading (101 or 19.7%) and 10 (1.9%) that turned out to be true, with one case of an unproven story.

Investigations which would yield 'unproven' results were avoided so as to not add to the confusion or be used to further one side of the argument. To address such claims, the team has turned to social media to run explainer videos which present all aspects in a measured, evidence-based manner, concluding with a call to action to verify first.

False

Misleading

True

Unproven

78.2%

19.2%

0.2%   1.9%
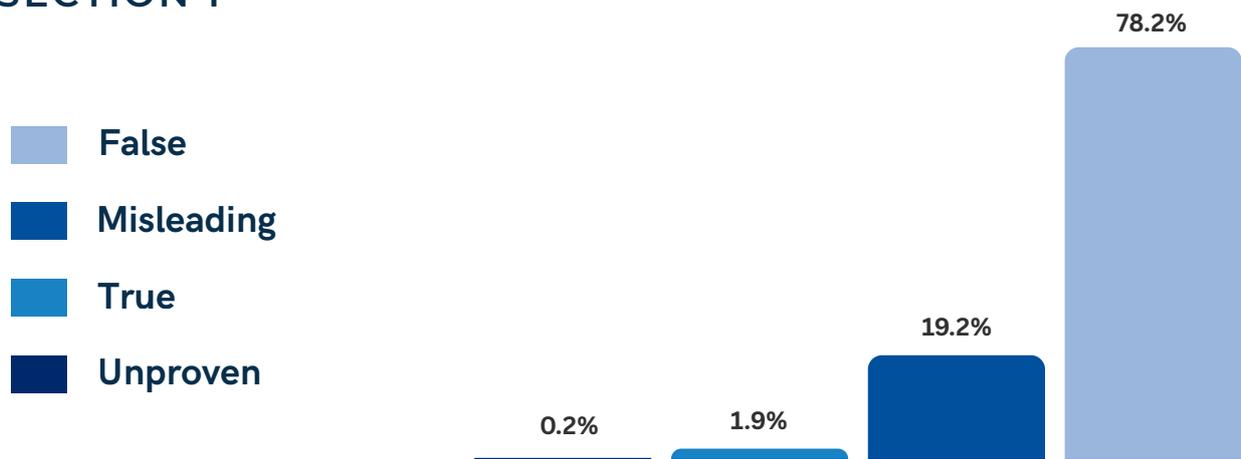
3. COMMON TOPICS USED IN MIS/DISINFORMATION

Of all the claims that iVerify catalogued from December 2023 to November 2025, 85.3% were political in nature, mostly relating to local politics and at times major international events, especially those concerning Islam, South Asians and high-interest news cycles.

The topics overlapping with politics were wide-ranging, including terrorism and militancy, pro/anti-military, diplomacy, social and rights movements, and anti-women.

Apart from sports and religion (see graph below), claims related to entertainment (2.2%) and climate (1.75%) were also noticeable.

85.3%

Politics

Climate

Entertainment

2.2%   1.75%
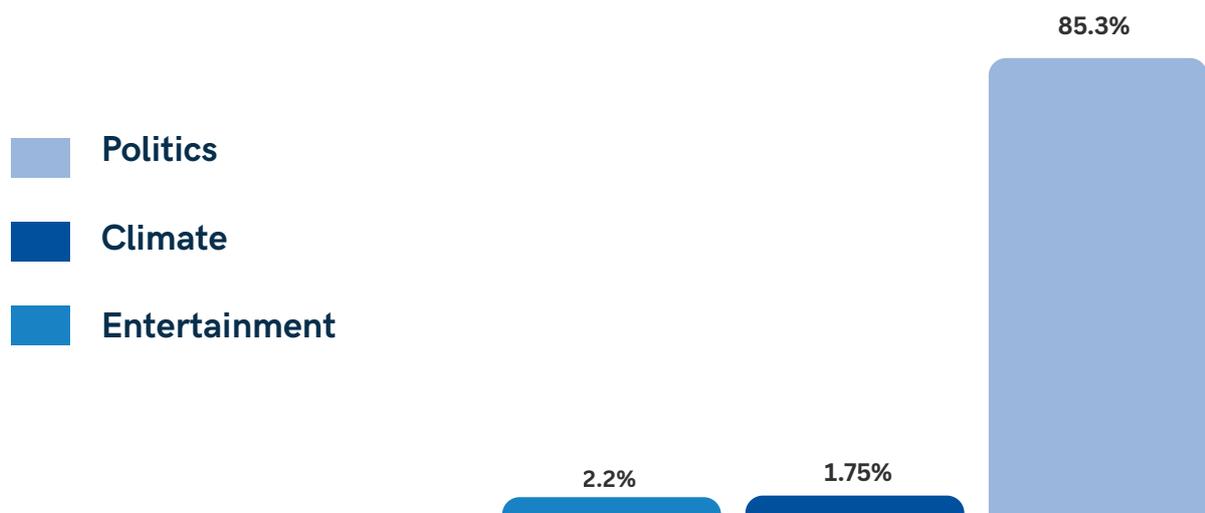
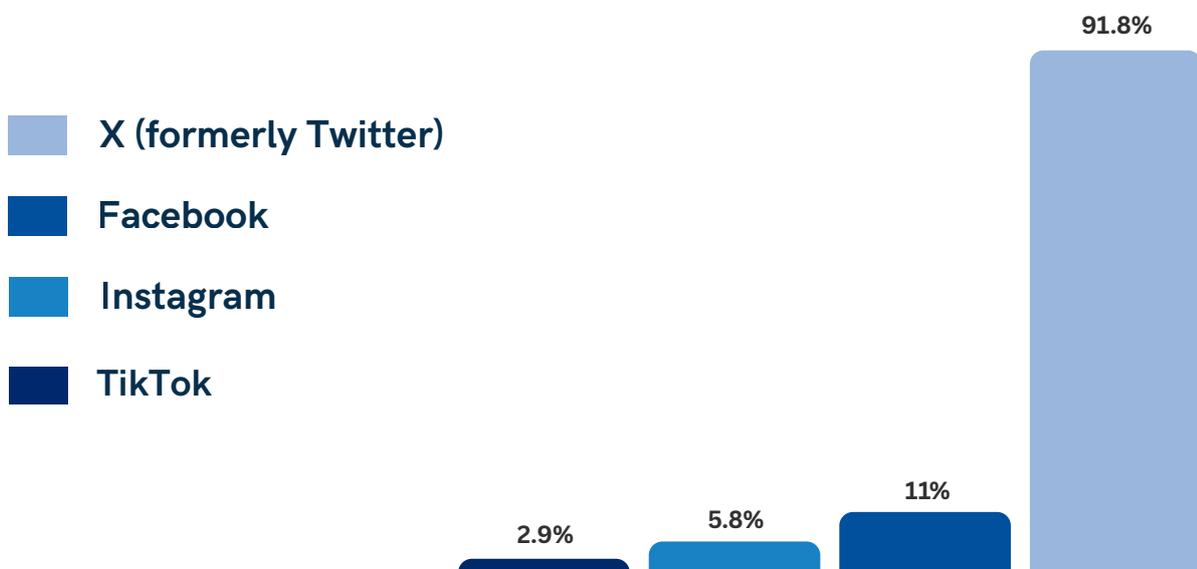## 4. USAGE OF SOCIAL MEDIA IN CIRCULATING CLAIMS

During iVerify's monitoring, social media platforms were found to be the most common medium to spread mis- and disinformation, with one claim often being shared across multiple platforms.

X was used in 91.8% of the claims, Facebook in 11%, Instagram in 5.8% and TikTok in 2.9%.

There is also an overlap in the data above as some claims were observed to be shared on multiple social platforms.

It is pertinent to mention that data cannot be collected on messages sent on WhatsApp, effectively taking it out of the running.

Throughout the 2024 election cycle and beyond, X emerged as the dominant platform for misinformation, highlighting its influence in shaping public discourse and amplifying misleading narratives – despite more than a year-long disruption[7] in its access which ended during the Pakistan-India escalation in 2025.

- X (formerly Twitter)
- Facebook
- Instagram
- TikTok

2.9%   5.8%   11%   91.8%

[7] https://www.undp.org/eurasia/dis/misinformation

5. ACTIVE MIS/DISINFORMERS

Based on the leads identified, PTI supporters* emerged as the most active group in spreading mis- and disinformation, present in 32.2% of the claims. This was followed by Indian accounts** (18.8%) and PML-N supporters (13.8%).

There is also an overlap in the data above as some claims were observed to be shared by multiple groups of perpetrators.

■ PTI supporters

■ Indian accounts

■ PML-N supporters

32.2%

18.8%

13.8%
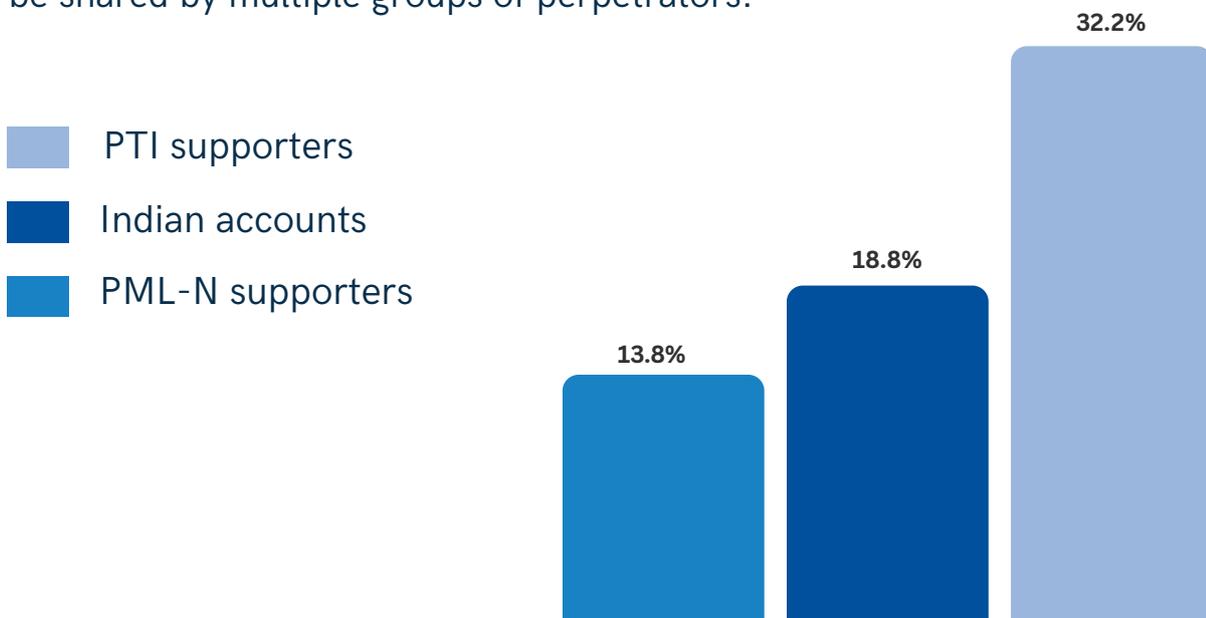
Collectively, these groups accounted for the vast majority of the identified perpetrators.

It is important to note that the claims identified by iVerify don't always take into account all the statements by government bodies or politicians that are shielded by national security or the element of confidentiality; news stories and journalists quoting unnamed sources; and WhatsApp users and forwards due to encrypted data.

Other groups that were observed spreading mis/disinformation include content creators, religious groups and their supporters, right-wing accounts, social media commentators, Afghan accounts and influencers.

*Supporters of political parties were determined by considering factors such as their posting history (retweets of party accounts and leaders, messages of support, etc) and visual information from their profiles (party logos, pictures of party leaders, etc).
**Indian accounts were a conglomerate of propaganda accounts and users, those sharing pro-Indian and anti-Pakistan material with the location on X set to India, right-wing accounts and journalists.

1.3 PATTERNS OF MIS- AND DISINFORMATION IDENTIFIED

During the course of these two years, iVerify identified some common patterns used to spread disinformation. The way the information is structured shows an intent to mislead people but once it has been shared, its dissemination is no longer contained to disinformation i.e. intentionally spreading false information.

- **Half-truth:** A piece of information that combines an element or part of the truth with false information. By first establishing a sense of credibility through the truth, it becomes easier to mislead – or at the very least confuse – the public.

Example: Social media posts circulating on January 10, 2024, alleged that the Canadian House of Commons[8] passed a resolution demanding the release of PTI founder Imran Khan. However, only a petition was presented on October 4, 2023, which did not demand Imran's release but did raise concerns about his arrest. This claim was manipulated by PTI supporters on platform X by premising it on a true development; it gained over 12,000 views.

- **Missing context:** Factually correct information, incidents or developments used without their original context. These are circulated at an opportune time, often when the news cycle is witnessing a situation of the same nature as the original context.

Example: On January 20, 2025, during Donald Trump's inauguration ceremony as the 47th President of the United States, PTI supporters began[9]

[8] https://www.iverifypakistan.com/news/1000010/no-resolution-was-passed-in-canadas-house-of-commons-demanding-imran-khans-release
[9] https://iverifypakistan.com/viral-video-of-us-president-donald-trump-praising-imran-khan-is-old-from-2023-not-linked-to-inauguration/

circulating an old video of him praising Imran Khan. Originally recorded in 2023, the video was falsely presented as a fresh statement linked to Trump's oath-taking ceremony. This repurposing of past content was an example of missing context, misleading audiences into believing that global leaders were making new endorsements aligned with local political developments.

- **Mistranslating:** Foreign languages, specifically their misinterpretation and incorrect translations, are used to mislead Pakistan's population that is either fluent in Urdu or regional languages. This method relies on the public believing the text's alignment with an accompanying visual or a common perception about an entity (person, country, organisation, etc) more than the text's correct translation.

Example: On day of the 2024 general elections, a video featuring *Sky News* correspondent Cordelia Lynch began circulating[10] on X, claiming she was reporting victory for the PTI on 150 seats. However, Lynch was delivering on-ground reports on the elections. Although she was speaking in English, the captions accompanying the video on several X posts were in Urdu, which translated to: "The international media has seen this news with their own eyes. Spread the word that PTI has won more than 150 seats." One of the posts had more than 333,000 views as the country was still awaiting results.

Sara Mir ✓
@SaraMirGilgity

International media has reported this news with their own eyes last night... that PTI has won more than 150 seats. Spread this news as much as possible.

Translate post

0:00 / 1:52

From **Cordelia Lynch**

[10] https://www.iverifypakistan.com/news/1000043

- **Impersonating accounts:** Fake accounts attributed to public figures are employed to mislead audiences, exploiting their trust in known public personalities, especially politicians and journalists, and institutions while taking advantage of fast social media scrolling habits as most people won't pay attention to the details revealing the giveaways of a well-made imposter account. Media accounts are also impersonated to run fabricated 'news reports', duping even those who would like to share only verified information.

Example: On December 11, 2024, claims regarding the death of Mumbai fugitive Dawood Ibrahim began circulating online and were widely shared by Indians. One such post was attributed to have been made by former caretaker prime minister Anwaarul Haq Kakar via his X account, a screenshot of which went viral. However, investigation revealed[11] Kakar made no such post and the said screenshot was that of a parody account.

- **Doctored/AI generated visuals:** From doctored government notifications to wholly AI-generated videos and deepfakes, manipulation of visuals with the aid of technical tools has been one of the most common ways to spread disinformation, and also malinformation (refer to Section 2).

[11] https://www.iverifypakistan.com/news/1000001
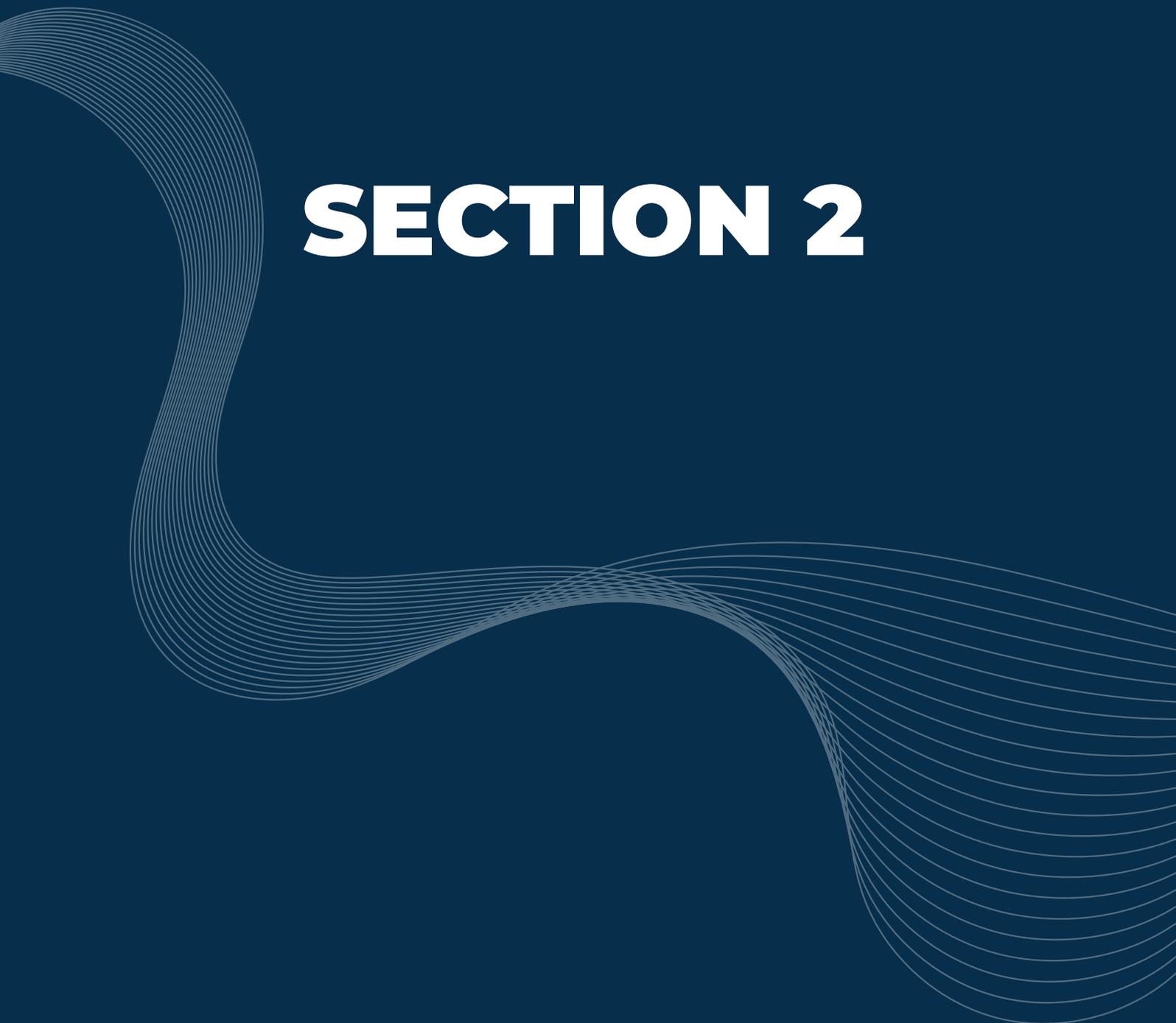
SECTION 2

AI AND MIS/DISINFORMATION

The capabilities of AI are both a bane and a boon. For all the creative possibilities that AI unlocks, it has become a tool for misinformation and disinformation on both domestic and global levels.

iVerify's interaction with AI-related mis- and disinformation started slow; there were relatively low numbers of leads in 2024 while 2025 saw a spike, mirroring real-world developments in this technology. Not only was AI used more widely, including as part of policy in organisations, but the rollout of more intuitive, user-friendly and advanced platforms, easily and freely available to the public, allowed more people to create realistic visuals – something the team could identify only during its course of investigation.

GROWTH IN AI-RELATED FACT CHECKS
(2024 VS 2025)
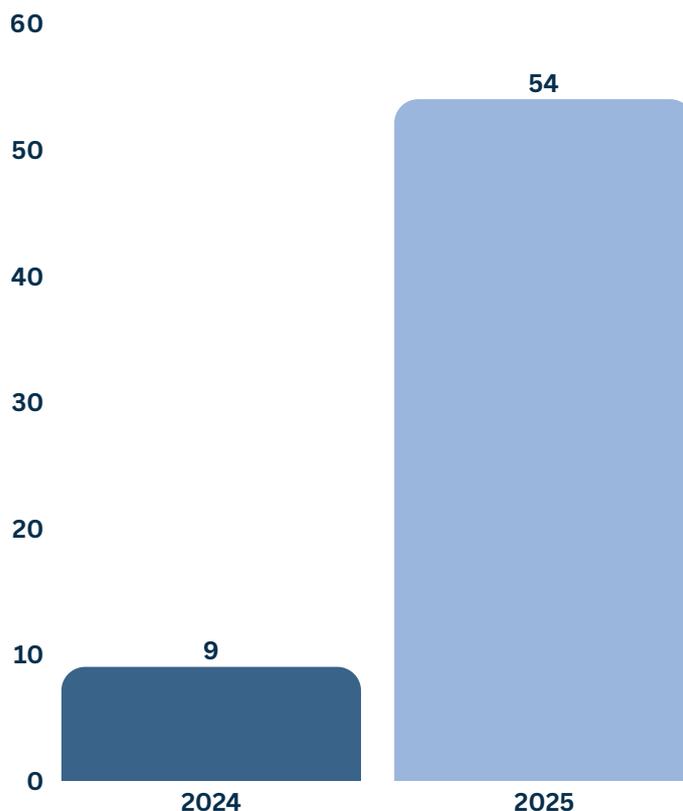
From nine fact-checks related to AI in 2024, iVerify witnessed a jump to 54 in 2025, demonstrating the increasing adoption of the technology in the mis- and disinformation landscape, particularly as it supplants older and cruder forms of image and video manipulation.

Furthermore, from initially being used to target politicians, particularly females, iVerify saw how AI-related mis- and disinformation soon encompassed various domains, including sports, entertainment, climate and business.

With more skilled AI-generated/assisted claims emerging in 2025, iVerify has had to constantly reevaluate fact-checking mechanisms and update them as advances in the technology are rolled out.

A dangerous trend spotted by iVerify is the "AI-fication of reality": when AI is used to make changes to otherwise original visuals, sometimes in the slightest but enough to mislead the public. Correspondingly, this showed a growing awareness in the public about the spread of AI content, with people increasingly asking chatbots to verify certain visuals or calling out deepfake content.

2.1 ELECTIONS - PRE AND DURING

In the pre-election period, PTI workers and supporters used advanced digital tactics to counter the government's attempts to exclude the party from the electoral process. In the absence of a big chunk of the party leadership, most of whom were in jail, the PTI used artificial intelligence to run its election campaign, i.e. deepfake videos, AI-generated voice clips, chatbots, and written content created to promote PTI's narrative. Allegations even surfaced[12] claiming that Imran Khan's piece for *The Economist* might have been AI-generated.

[12] https://www.dawn.com/news/1804477

SECTION 2

When compared to other countries that went to polls the same year, the advanced use of AI in spreading mis- and disinformation leading up to elections in Pakistan remained low.

iVerify found some claims which involved images that had been tampered with, but none with a highly technical use of AI. On January 30, less than two weeks before the elections, when Khan and his close aide Shah Mahmood Qureshi were convicted in the cipher case, a manipulated image of the PTI founder sporting a beard began circulating online. While the public's interest in the PTI founder's appearance following his incarceration was high, online discourse showed that skepticism around the photo was also high. Investigation confirmed[13] that the image was indeed doctored, with Fake Image Detector saying it was either computer-generated or modified.

Viral Image of Former Prime Minister Imran Khan

[13] http://iverifypakistan.com/news/1000030

SECTION 2

The most dangerous trend iVerify Pakistan spotted was the intention to interfere in the electoral process by confusing or misguiding voters. Case in point: a purported audio of Khan announcing a boycott of the elections circulating online. It was, however, immediately debunked by the PTI. Similar attempts, although not necessarily using AI, were made via religiously charged posts against political opponents, false claims, frequent recurrence of misinformation regarding the withdrawal of candidates from the electoral race, and other misleading claims.

In contrast, during the period after elections, PTI's use of AI technologies saw a decline as the party's focus moved towards challenging alleged discrepancies in the election results. But that did not mean that AI was forgotten.

In July, PTI supporters appeared to leverage AI technology to create an endorsement from Bill Gates, generating a deepfake video[14] of the former Microsoft CEO allegedly praising Imran. The original post featuring the video gained over 157,000 views on X.

Using a host of tools — reverse image search and voice comparison techniques — iVerify Pakistan investigated the authenticity of the video and found that Gates' voice in the clip was that of an AI-generated clone. Meanwhile, the video was repurposed from an old interview.

14 https://www.iverifypakistan.com/news/1000202

2.2 POLITICAL TARGETING

Pakistan is a politically charged country and politics dominate most public discussions; leads related to AI were not immune to this trend either. iVerify observed that AI-generated content and visuals on political developments and politicians were primarily spread by party supporters to bolster and prop up their leaders while spreading propaganda to undermine and diminish opponents.

The common players were supporters of the PTI and the PML-N. In the PTI's case, the party's campaigners used key international events such as the US elections to build on their narratives using AI, i.e. the dissemination of doctored videos of US President Donald Trump[15] and his aide Richard Grenell[16].
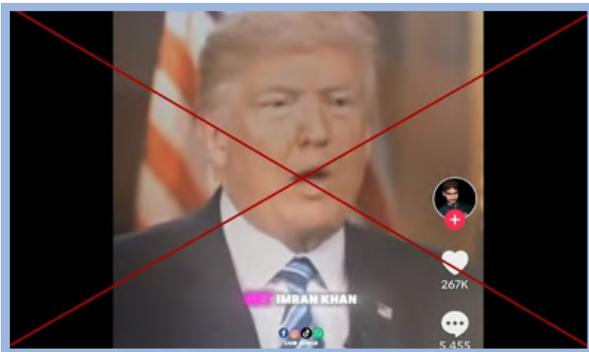
The videos were used to lend the impression that they were voicing support for Khan. Conversely, for the other side of the camp — the ruling PML-N — Khan was a key target. iVerify noted multiple instances of the ex-premier's old videos being manipulated to push an anti-PTI narrative.

In one example, a purported video was circulated online wherein Imran Khan could be seen kissing two women. An analysis[17] of the screenshots from the video showed the same distortions common to deepfakes.

[15] https://www.iverifypakistan.com/news/1000212
[16] https://www.iverifypakistan.com/news/1000248
[17] https://www.iverifypakistan.com/news/1000231

SECTION 2

Across both sides of PTI and PML-N supporters, the use of AI-generated visuals to spread scandalous content about female politicians was another common observation. Even though the crude nature of such content was at times observed to hold low credibility with people questioning their authenticity, the nature of such claims makes them dangerous even with low believability (refer to 2.4).

Beyond rivalries, iVerify observed another trend wherein misinformation on political events was spread using exaggerated AI imagery. Generative technology was also used to push a particular narrative by dubbing videos of key government representatives or major politicians.

In March 2025, Prime Minister Shehbaz Sharif became a target of this trend when a viral video allegedly showed him being confronted over funding Umrah for a 30-member government delegation. When investigated, it was found that the clip was dubbed over[18].

Similarly, a video of PPP Chairman Bilawal Bhutto-Zardari was shared by Indian accounts in July 2025 that allegedly showed him saying that people affiliated with Pakistan were present in India's parliament. An investigation showed the video was doctored[19] and he said no such thing.

[18] https://www.geo.tv/latest/597378-fact-check-doctored-video-falsely-claims-saudi-anchor-criticised-pm-shehbaz-sharif-over-umrah-expenses
[19] https://www.iverifypakistan.com/news/1000389

2.3 WARS AND CONFLICTS

During its operations, iVerify Pakistan monitored several periods of political instability and conflicts, but the issue became a running theme in 2025. The major conflicts covered in the last year included the India-Pakistan standoff in May, the Iran-Israel war in June, and the breakout of border skirmishes between Afghanistan and Pakistan in October.

For the Pakistan-India tensions, the team logged 11 leads since the terrorist attack in Indian-held Kashmir's Pahalgam on April 22, 2025. All of them were pursued, In May, 33 leads were logged related to the conflict, of which 18 were pursued. During the course of the Iran-Israel conflict, iVerify identified at least 38 claims circulating on social media and investigated 32 of them.

Periods of conflict proved to be some of the most intense times for fact-checking operations, with a high daily influx of leads and claims. Overwhelmingly, the leads pursued were false instead of misleading, which indicates there was a deliberate intent to peddle misleading fake information instead of merely sowing confusion.

In today's world, wars are no longer just fought on battlefields; they have now moved to the reel world, where truth has become both a casualty and a weapon. In all of the above three instances, iVerify closely monitored how AI was weaponised during active battles and periods of friction.

2.3.1 SHAPING PUBLIC PERCEPTION

A key role that AI-generated content played during these conflicts, as observed by iVerify, was in shaping and moulding public perception for both domestic audiences and the opposing side.

During such high-intensity periods, information is fast and immense, making social media users vulnerable to getting lost in minute-by-minute updates. The spread of doctored content, therefore, creates doubt, uncertainty and confusion.

One particular example was a deepfake video[20] of Prime Minister Shehbaz during the Pakistan-India conflict, which was tampered to show him conceding defeat to New Delhi. "The altered version used AI voice cloning and lip-sync technology to deliver a narrative that was entirely fabricated – but disturbingly believable," the International Federation of Journalists observed[21] in a blog.

Another instance was an altered video of the military spokesperson, wherein he allegedly claimed[22] the loss of two Pakistan Air Force fighter jets to the neighbouring country.

[20] https://www.sochfactcheck.com/footage-of-shehbaz-sharif-accepting-pakistan-defeat-to-india-is-ai-generated-war-2025/

[21] https://www.ifj.org/media-centre/blog/detail/category/ai/article/ifjblog-ai-deepfakes-and-the-fog-of-war-disinformation-in-the-2025-india-pakistan-conflict

[22] https://factcheck.afp.com/doc.afp.com.46GJ7NP

SECTION 2

The intonation and body movements in the military spokesperson's video were difficult to distinguish from his real press conferences. The video was highly realistic with almost perfect lip-syncing, which can be attributed to improvements in technology that make facial and vocal edits appear almost seamless. However, analysing the audio of the footage and comparing it with other videos of the military spokesperson showed a monotonous and robot-like tone. Deepfake detectors such as Hive Moderation and Attestive showed results of the audio likely to be 99% AI-generated and a 91% suspicion rating, respectively.

The purpose of the deepfake was clear: to fool the audience into believing the video by using cues that were difficult to detect. However, they did not hold up.

During the four-day conflict and after, misleading AI visuals circulated on the other side of the border as well, i.e. deepfake videos[23] of Indian Prime Minister Narendra Modi and Home Minister Amit Shah apologising to Pakistan.

The observation underlined that such deepfakes can often be distinguished from reality if investigated thoroughly, but that is not the purpose; the aim is to just look and sound believable enough at first glance. Especially in times of conflict and war, when information is dumped on the internet at an overwhelming speed and quantity, it becomes easier for users to believe in any and all multimedia posted online, as the line between fact and fiction blurs.

At the same time, there is a lack of access for the media in such circumstances; hence, the public can only consume official versions from both sides, which leads to higher believability in what is being spread through public discourse. In the case of Pakistan and India, people from both sides were on the lookout for what each country was saying – beyond official statements – adding to the virality of unverified information. Closer coordination between the government and fact-checkers could assist in this misinformation whirlpool.

The trend continued even after the conflict ended, with iVerify fact-checking

[23] https://reutersinstitute.politics.ox.ac.uk/news/truth-casualty-how-indian-fact-checkers-debunked-false-claims-during-india-pakistan-crisis

five doctored videos of Indian officials and celebrities criticising their authorities and propagating Pakistan's narrative regarding loss and damage in the conflict, proving that AI content not just contributes to wartime propaganda but is also used to achieve post-conflict goals of creating public perceptions.

A similar but more intense trend was once again seen during the Iran-Israel conflict, where eight out of a total of 32 leads fell into the category of AI content and deepfakes. One such video shows[24] Iranian soldiers thanking the Pakistani intelligence.

It gained over 1.4 million views and was widely shared on social media.

Another example was a video of President Trump, in which he called on Pakistan to remain neutral during the Iran-Israel war. Shared by a PML-N supporter and viewed over 620,000 times, the clip was confirmed[25] to be a deepfake.

The trend carried forward in October's Pakistan-Afghanistan skirmishes, when iVerify detected a video of Russian President Vladimir Putin warning Islamabad against attacking Kabul. A day later, the Russian embassy debunked[26] the video as fake.

[24] https://www.iverifypakistan.com/news/1000357
[25] https://www.iverifypakistan.com/news/1000339
[26] https://x.com/RusEmbPakistan/status/1983795057926009153

2.3.2 HIGH VOLUME, HIGH SPEED

Old visuals formed a bulk of the material that was used in the mis- and disinformation cycles during wars. A bulk of fake images and videos spread during the Pakistan-India conflict were recycled visuals from the Gaza invasion, Israel's conflicts with Iran, and even previous altercations between Pakistan and India. The visuals were shared in low quality to obscure telltale signs that could prove they were unlinked to the event. The International Federation of Journalists also noted: "Old techniques remained just as potent. Several widely shared videos were not AI-generated but were instead repurposed from past or unrelated events."

When the purpose is to spread propaganda during times of conflict, the quantity of such visuals matters more than their authenticity. One such example was a video that was widely circulated as showing an attack on Pakistan by India, racking up over 5.1 million views. However, the clip had nothing to do with the conflict and instead showed[27] Iranian missiles targeting the Nevatim airbase in Israel in October 2024.

Old imagery from the Gaza conflict was, in particular, a common go-to source for false and misleading visuals that were shared during the conflict, as another video[28], said to be of explosions and airstrikes in Pakistan, racked up 1.6m views. However, it was actually Israel's bombing in Gaza in October 2023.

The aim behind recycling these visuals is simple: exaggerate the scale of attacks, spread misleading facts, and inflame public sentiment. The aforementioned Nevatim airbase attack video was also shared during the first day of the Iran-Israel conflict, to craft the narrative that Iran had already suffered significant attacks early in the war.

Other old methods observed during the two conflicts, as well as others, were the sharing of material and visuals without their original context, doctored visuals through traditional techniques such as stitching two separate videos, editing an image, and impersonating official accounts.

Several posts during the Iran-Israel conflict involved the impersonation of official government statements, a tactic designed to fabricate support or opposition from specific countries. They were mostly shared by partisan or ideologically motivated accounts and sought to provoke geopolitical friction. The iVerify team identified three leads of impersonation.

On June 16, 2025, an Iranian news outlet, *Iran in Arabic*, on its X account, claimed that Afghanistan had issued an official warning against the US against military intervention in Iran. The fabricated[29] statement was used to position Afghanistan as taking a bold anti-US stance, manufacturing the illusion of regional alignment against American involvement.

---

28 https://www.iverifypakistan.com/news/1000320
29 https://www.iverifypakistan.com/news/1000344

SECTION 2

Just like the aforementioned instance, iVerify Pakistan observed how even journalists and media outlets were susceptible to spreading mis- and disinformation during times of conflict because of the rush of content and news.

The problem was particularly egregious in the Pakistan-India conflict. On the other side of the border, primetime news channels aired exaggerated information, with hysteria reaching a point where the Indian government accounts had to step in and advise[30] media outlets against senseless reporting. Although not the same in volume, Pakistani media outlets and journalists too shared misleading and fake content.

On May 10, senior journalist Hamid Mir shared a video of a fire brigade approaching a site clouded with thick smoke, as onlookers observe and guide the firefighters towards the scene. He said the video, which gained 1.7 million views, showed the Udhampur air base. However, the clip was actually of a factory fire in India from May 9, a day before the military escalation.

Even during the Iran-Israel conflict — a landmark case[31] in AI-driven disinformation — a senior journalist affiliated with *Middle East Eye* shared a doctored clip of Mohsen Rezaei, a member of Iran's National Security Council, falsely claiming[32] that Pakistan had assured Iran it would retaliate with a nuclear strike if Israel dropped a nuclear bomb on Tehran.

The misleading clip, viewed by over 900,000 users, was widely amplified, including by several Indian media personalities. In reality, Rezaei made no such statement.

[30] https://www.washingtonpost.com/world/2025/06/04/india-news-channels-misinformation-pakistan-conflict/
[31] https://carnegieendowment.org/research/2025/07/iran-israel-ai-war-propaganda-is-a-warning-to-the-world?lang=en
[32] https://www.iverifypakistan.com/news/1000341

SECTION 2

One pattern that iVerify observed during the India-Pakistan conflict was the proliferation of outlandish claims — which would normally not merit even serious discussion — broadcast as breaking news on Indian media.

From the Indian side, these claims — such as the Karachi port attack[33], the invasion of Islamabad and a coup against the army chief — were spread in tandem with the timing of the actual attacks on May 9.

The tactic served multiple purposes: exaggerate the actual scale and success of the attacks to cause confusion, serve to cover, hide and take attention away from actual military strikes, sustain hysteria among the Indian population, and provide content to feed the propaganda machine.

The believability of such claims was not high despite their high virality, as in the case of the Karachi port attack, where one post gained 2.9m views and was broadcast by Indian media outlet *Zee News* as well. In this instance, the believability among Indian users too was low, with most even calling for restraint.

[33] http://iverifypakistan.com/video-does-not-show-karachi-port-destroyed-by-indian-attack-but-plane-crash-in-philadelphia/

2.4 ANTI-WOMEN

Pakistan is regularly placed at the bottom of international indices related to women's rights, appearing at 148 out of 148 countries in the Global Gender Gap Report[34] in 2025. Enter a host of powerful and unsupervised AI tools, functional on existing gender biases, now enabling that violence to proliferate in complex ways. During its two-year operations, iVerify identified a trend where AI was used to create new forms of abuse and amplify existing ones, using women as softer targets, especially those under the public eye.

Non-AI
5

Total
11

AI-Generated
6
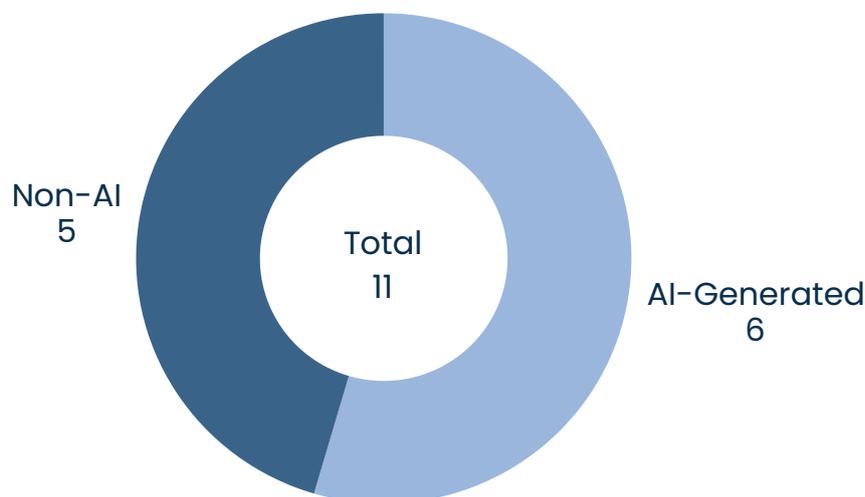
Of 11 pursued fact-checks that were deemed to be "anti-women", iVerify logged the use of AI in six of them to target women through doctored visuals, including politicians and celebrities.

The purpose is clear: to undermine their credibility, a tactic that is increasingly being used by political players against women of the rival camps. The way it is done follows a well-thought-out process, which involves using AI to generate scandalous videos and photos. These clips are then disseminated from not just a few but dozens of accounts across social media, each gaining views in hundreds, thousands, even millions at times.

Punjab Minister for Information and Culture Azma Bokhari became a victim of this new form of abuse in 2024, when a sexualised deepfake video[35] began circulating. And despite being in an important position in a key province, the online harassment and character assassination that came her way did not fall short.

[34] https://www.dawn.com/news/1916743
[35] https://www.iverifypakistan.com/news/1000137

SECTION 2

Punjab Chief Minister Maryam Nawaz and former first lady Bushra Bibi are the other two common targets, whose fake videos and images pop up on social media every time the political parties affiliated with them lock horns.

This online harassment is not just restricted to the political sphere; in 2025, journalist Benazir Shah was targeted when a deepfake video[36] claimed to show her dancing inside a nightclub.

Earlier, iVerify noted a trend of manipulated videos showing mainstream Pakistani actresses in intimate scenes — all of this content was entirely fake.

Other female celebrities were also targeted in the same manner, such as scandalous imagery such as concocted scenes of them kissing male actors or unrelated videos or photos meant to shame them on their dressing.
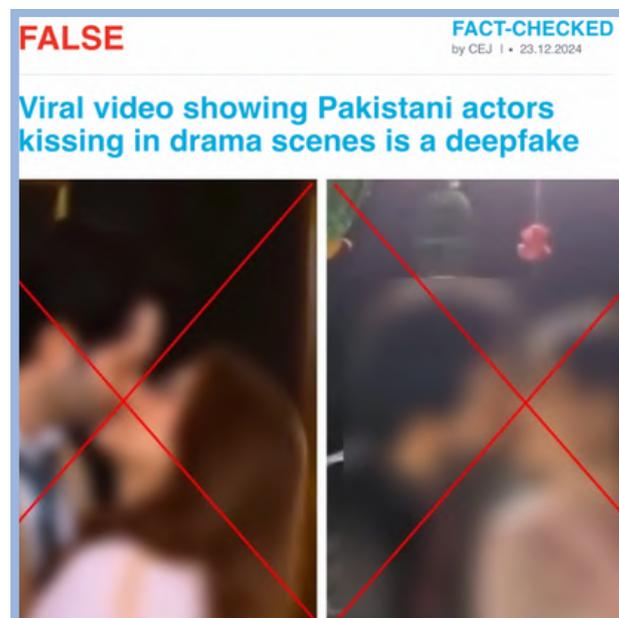
These campaigns not only inflict emotional distress and reputational harm upon the targeted individuals but also perpetuate a culture of toxicity and misogyny within public discourse. It serves to intimidate and silence voices, particularly those of women, thereby impeding their ability to participate fully in public life and contribute to meaningful dialogue.

---

[36] https://www.dawn.com/news/1956689

2.5 AI-FICATION OF REALITY

During its two years of operation, iVerify Pakistan identified an alarming trend: visuals from real-life incidents being amplified to exaggerated levels. These visuals are often presented in ways that make audiences likely to believe events that may not have occurred at all.

A recurring example of this phenomenon can be seen in AI-generated content featuring former prime minister Imran Khan. Such visuals, widely circulated online, depict him doing push-ups[37] in Adiala Jail, jogging on a treadmill, or reading inside a prison cell. Despite having been incarcerated for over two years, these images construct a version of reality that does not exist — yet remains plausible enough to be believed by many.

While the intent behind such content may appear harmless in some cases — framed as insider access or unseen footage — the practice raises serious ethical concerns. It can diminish the gravity of real events or inflate public sentiment around incidents that never took place. The unverified nature of AI-generated visuals creates confusion, dilutes the impact of genuine situations, and makes it harder for audiences to focus on the underlying issues.

FALSE   FACT-CHECKED by CEJ | 15.12.2025

Viral image of Imran Khan doing push-ups in Adiala jail is AI-generated

One such instance occurred in 2024, when journalist Hamid Mir shared a photograph of activist Romasa Jami allegedly being thrown to the ground during a clash with law enforcement at the Sindh Rawadari March outside the Karachi Press Club. The post, accompanied by a caption condemning police brutality, was widely circulated. However, closer scrutiny revealed that the image was AI-generated, displaying multiple visual inconsistencies. Other users on X later shared original footage showing Jami being dragged along the road by police during the same protest.

---

37 https://www.iverifypakistan.com/news/1000512/viral-image-of-imran-khan-doing-push-ups-in-adiala-jail-is-ai-generated

SECTION 2

While Mir's actions may have been motivated by an intention to highlight police violence, the ethical implications of using AI-generated imagery to depict real-life events remain significant and cannot be overlooked.

This trend also spilled into conflicts, where AI-generated content was used to spread images and videos that allegedly depicted wartime scenes and visuals, whereas investigations showed otherwise. One widely shared post during the 2025 Iran-Israel war featured[38] an AI-generated image of mid-rise buildings collapsing in Israel, claimed to be the result of Iranian missile strikes. In reality, though, the clip was created using synthetic imagery.

The video aimed to showcase and valorise the Iranian military's performance while showcasing the failure of Israeli defences. The same use of AI imagery to spread an exaggerated version of reality was also seen[39] in the Pakistan-Afghanistan conflict later the same year.

38 https://www.iverifypakistan.com/news/1000358
39 https://www.iverifypakistan.com/news/1000450/viral-image-of-alleged-afghan-taliban-attack-on-pakistans-torkham-border-crossing-is-fake

2.6 2026: WHAT'S AHEAD FOR AI AND DISINFORMATION?

"We're going to move from assuming what we see is real by default, to starting with scepticism. Paying attention to who is sharing something and why."

This insight was among Adam Mosseri's, the head of Instagram, reflections[40] on the last day of 2025. iVerify observed this trend increasingly in the second half of the year; as the use of AI became more and more rampant, a parallel rising trajectory — while not at the same speed — of the public paying attention to the content and flagging it as "fake" also became more and more apparent. At the very least, iVerify noted that people are questioning, more so than before, whether a visual is authentic or fabricated.

"AI trust paradox" is a phenomenon[41] global researchers have already identified; as AI outputs become more persuasive and audiences struggle to differentiate fact from fabrication, trust in both news media and digital platforms will erode further. At the same time, this increases the possibility that trusted media and platforms will be sought out by the public once again as the need to know what is real, amidst the inundation by AI slop, outweighs the trust deficit in mainstream media. With consumption of information continuously shifting towards digital[42], media outlets will have a bigger role to play in claiming space on social media.

For fact-checkers, misleading information and visuals will require the most attention, as they have the potential to deceive even those looking for evidence of AI-tampering. More so, fact-checkers may have to pivot towards flagging authentic content more than verifying false claims. But the role of fact-checking itself will need to be defined and structured through participation of credible media houses, especially in light of large social media platforms like Meta and X disbanding[43] fact-checking initiatives.

[40] https://www.instagram.com/p/DS7pz7-DuZG/?hl=en&img_index=1
[41] https://www.kantar.com/north-america/inspiration/ai/navigating-the-ai-trust-paradox
[42] https://gallup.com.pk/post/39135
[43] https://www.wired.com/story/meta-ditches-fact-checkers-in-favor-of-x-style-community-notes/

SECTION 2

The *BBC* launched its Verify service in May 2023, going a step further in June 2025 with Verify Live[44] to tell audiences in real time what claims they are investigating and how they are being checked – such integration of fact-checking within media models will have to be explored more actively in 2026 to counter the 'Community Notes' model that allows users to publicly flag content they believe to be incorrect or misleading. In Pakistan, fact-check accounts[45] cropping up on X, that lean more towards labeling what is true or false – or sometimes outright propaganda[46] – instead of presenting concrete evidence, are further weakening infrastructures on social media that support truth and accountability.

Experts highlight[47] that AI-generated deepfakes, voice cloning, and synthetic images are approaching near-indistinguishability from authentic content; this raises the question whether social media platforms will monitor and regulate AI labels aggressively, and how. And if they don't – a reasonable possibility given their actions regarding fact-checking models – then the abovementioned hybrid models of media and fact-checkers become ever more critical.

We're still in the early days of AI but how we act on the emerging patterns will help shape its future, especially as industries race to deploy AI agents into their workflows, less funding and support are allocated for verification mechanisms, and the growing trend of state-sponsored fact-checking initiatives[48] devalue legitimate fact-checking efforts.

[44] https://www.bbc.co.uk/mediacentre/articles/2025/bbc-verify-live
[45] https://x.com/fcheckmaster
[46] https://x.com/azaadfactcheck?lang=en
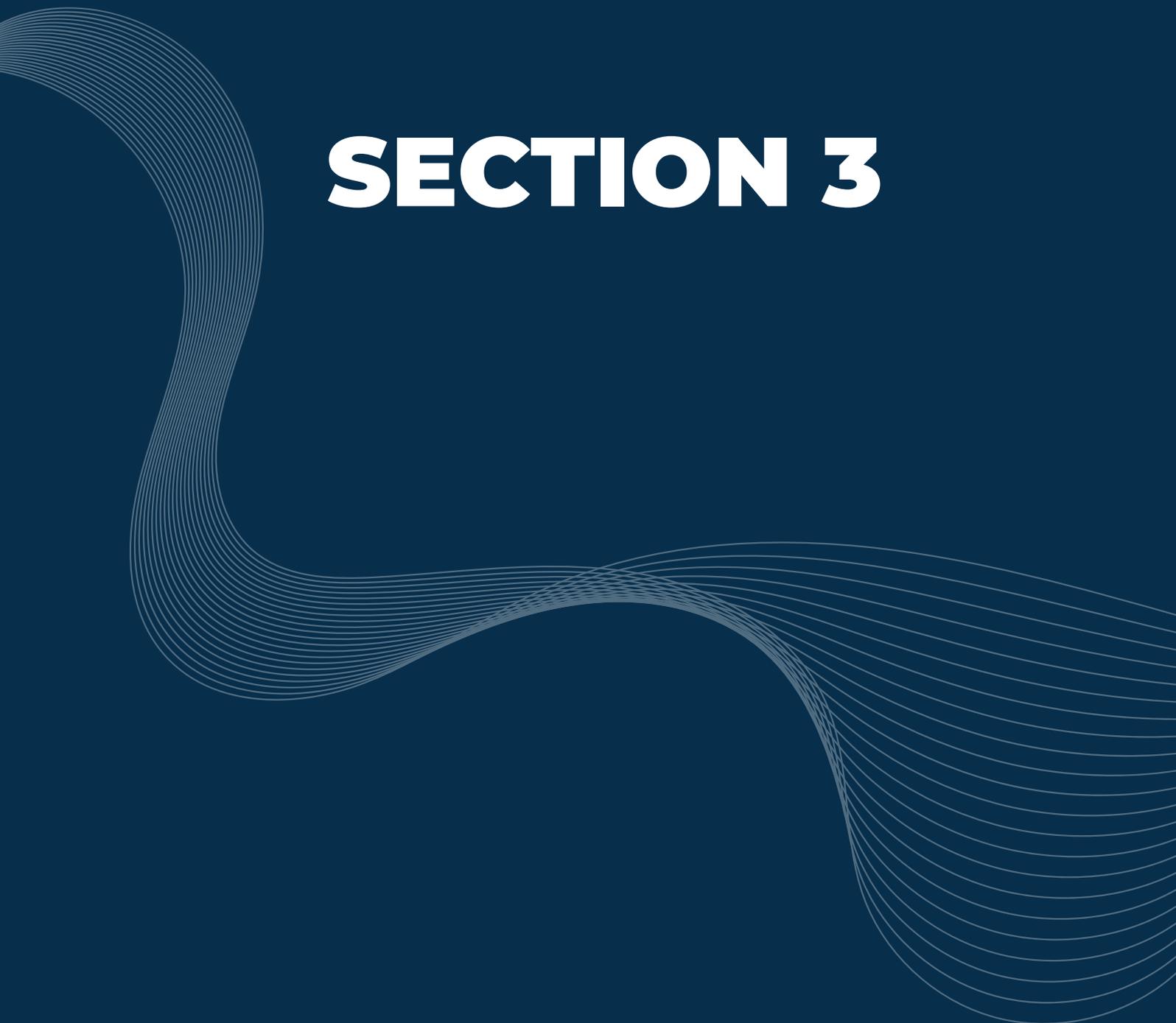[47] https://www.forbes.com/councils/forbesbusinesscouncil/2025/12/30/the-year-authenticity-breaks-deepfakes-ai-and-the-battle-for-trust-in-2026/
[48] https://www.bbc.co.uk/rd/articles/2025-09-misinformation-disinformation

SECTION 3

POLITICAL TARGETING AND HATE SPEECH

One of the cornerstones of iVerify is to monitor harmful content and the team, throughout its operations, has noted its constant presence on the internet, manifested in different ways, including gender, ethnicity, religion and political affiliation.

Harmful content encompasses rude, disrespectful, or unreasonable comment/behaviour around issues that contain toxicity, obscenity, threats, insults, and hate speech – the trickiest of all to tackle. An argument can be made that all unverified information is harmful in some capacity to an individual or community, or even a cause. iVerify Pakistan analysed its data to identify what kind of claims caused harm on a large scale and in the majority cases, they traced back to the news cycle and public sentiments – both reflected in widespread circulation of similar posts, multiple claims around the same topic and a persistent push in peddling that information.

It is worth addressing that while we can't establish intent of an individual in spreading harmful information (on purpose or not), we analyse how it is being spread, the topic at hand with its historical context, study different actors involved in spreading the information and other factors to reach our conclusions. This has been explored further in Section 3.2.

A common element observed in harmful content centred on ethnic and sectarian issues is the depiction of targets as anti-state. Painting certain groups as subversive through mis- and disinformation reduces public support for them, invites official action and justifies any subsequent state clampdown. Labelling ethnic or sectarian groups 'anti-state' when they speak up for their rights sucks the oxygen from the validity of their demands and casts them as traitors.

3.1 PUBLIC SENTIMENTS AND THE NEWS CYCLE

Across varying news cycles, the targets of harmful content changed – but more or less, as evident in the data – they remained connected to politics.

Here, iVerify provides a snapshot of how news waves brought with them an influx of mis- and disinformation against:

POLITICAL PARTIES

iVerify Pakistan began its operations just before the 2024 general elections. It was a time when the PTI was in disarray, with its top brass jailed and other party leaders 'sidelined' by all players, especially the then caretaker government. Consequently, frustration boiling inside party supporters spilled online, reflecting in the types of claims and leads the team encountered.

The caretaker setup and members from the opposite side of the camp — the Pakistan Democratic Movement — became key targets. Stories and leads in election week were subsequently focused on political participation, with two fact-checks investigating the PML-N boycotting[49] the polls and senior party politician Khawaja Asif quitting politics. From the other end, PML-N supporters used the PTI's lack of leadership to create confusion among its voters. Leads from PTI supporters were seen to spread content about foreign media and actors confirming the party's victory in the polls.

Curiosity over election results was the subject of immense public interest, but it grew after the Election Commission of Pakistan delayed their announcement, leading to widespread allegations of rigging by the PTI and other political parties.

[49] https://www.iverifypakistan.com/news/1000040

Since the polls and the subsequent formation of the coalition government, iVerify observed that the Prime Minister Shehbaz Sharif became a frequent target of mis- and disinformation both locally (the opposition) and across the border. The team noted a significant involvement of Indian accounts during periods of the premier's foreign visits and engagements i.e. his trip to China[50] in September 2025 and his address at the United Nations General Assembly.

In the first instance, he was claimed to have been sidelined in the group photograph of world leaders, while his speech[51] and actions[52] at the UNGA were misrepresented — both aimed at not just undermining his stature in the global community but also harming Pakistan's international image using either doctored or manipulated content.

JUDICIARY

Public interest in the affairs of the judiciary remained high during the two-year period, particularly in former chief justice of Pakistan Qazi Faez Isa for two reasons: his verdicts in the PTI election symbol case and the Mubarak Sani case involving the Ahmadi community.

[50] https://www.iverifypakistan.com/news/1000413
[51] https://www.iverifypakistan.com/news/1000435
[52] https://www.iverifypakistan.com/news/1000437

SECTION 3

On January 13, 2024, Justice Isa deprived the PTI of its electoral symbol — a cricket bat — after marathon hearings, and henceforth became a target of several online attacks and criticism by party supporters. The digital discourse was, per iVerify's findings, also reflective of a wider sentiment among other social actors. Interest surrounding him continued throughout the first year of iVerify's operations, with a doctored video[53] shared all the way later in March 2025 in which he could be heard shifting the burden for verdicts against the PTI toward the system.

In February 2024, his verdict in a case involving an Ahmadi man and ordering his release earned the ire of religious quarters, with Jamiat Ulema-i-Islam Fazl (JUI-F) chief Maulana Fazlur Rehman sharing a video that accused the judge of allowing the minority community to freely propagate its beliefs. This intense pressure led to a subsequent review of the verdict.

Both the controversies surrounding him culminated into a unique incident in September 2024, when PTI supporters widely shared a video purportedly showing a doughnut shop employee misbehaving with Justice Isa due to public frustration over his verdict against the PTI and its role in the party's electoral fortunes. The video was accompanied by a claim that the employee's family had been abducted and the shop was sealed.

However, iVerify's investigation[54] showed that not only was the picture of the sealed branch different from the actual location the judge visited, but the video was also of a prior altercation in August 2024.

[53] https://www.iverifypakistan.com/news/1000278
[54] https://www.iverifypakistan.com/news/1000189

Ongoing developments in the news cycle concerning other judges also impacted how mis- and disinformation was spread against them, with a deeper dive later on how coordinated campaigns were run against some of the six Islamabad High Court judges who flagged the alleged involvement of intelligence agencies in judicial affairs.

iVerify observed that this interest and momentum in mis- and disinformation about judges, particularly those of the Supreme Court, remained high in 2024, with speculation over who would be the next chief justice and a crucial verdict in favour of the PTI in the reserved seats case in July 2024.

However, iVerify also saw that this interest waned over time with the passing of the 26th Constitutional Amendment, that also saw SC's Justice Mansoor Ali Shah being stepped over as the chief justice in October 2024, and the passage of the 27th Constitutional Amendment in November 2025 — developments that were described as the "death" and "funeral" of judicial independence by the legal community.

With a widely perceived decline in the importance of the judiciary, which also impacted the flow of mis- and disinformation regarding judges and judicial affairs, as a substantial decrease was noted in claims and leads regarding them outside of any fresh developments in the news cycle.

ARMY

A similar high level of public interest was observed in affairs related to the military and army chief as well, both on a domestic and international level, which were majorly impacted by the institution's relations with the PTI, and in the aftermath of the conflict with India in May 2025.

When the military spokesperson criticised the PTI in a lengthy press conference in May 2024, he became a target of vitriol and criticism online. iVerify Pakistan logged seven leads the same year, all related to the army chief. Six of these were pursued. The number jumped to 42 by November 2025, of which 16 leads were pursued.

The surge was linked to the military and the army chief's growing prominence in international affairs in 2025, particularly the boost to Pakistan's image following the conflict with New Delhi.

Subsequent developments — his appointment as field marshal, high-profile meetings with Donald Trump, thaw in US-Pakistan relations, the signing of a defence deal with Saudi Arabia and Pakistan's role in the Gaza ceasefire agreement — also impacted this trend, with the Pakistan Army and its chief gaining greater prominence and interest from the domestic and global public.

ETHNIC GROUPS

During this time, Balochistan's issues and long-simmering tensions between the PTI-led Khyber Pakhtunkhwa and PML-N-led Punjab provinces also reflected public sentiment along ethnic lines.

In the aftermath of the PTI's Islamabad protest in November 2024, a state crackdown led to the police racially profiling and victimising Pakhtuns. This also translated online, where hate speech was used to cast Pakhtuns as having a violent and disruptive mindset, used to justify the alleged actions of state authorities against the ethnic group.

PML-N supporters[55] shared a video on various social media platforms on December 16, 2024, that featured several men fighting in the middle of a road, with users claiming that it highlighted increasing lawlessness in KP. However, the video was actually of an old road rage incident from 2016.

MISLEADING

FACT-CHECKED
by CEJ | • 19.12.2024

Viral video of KP road rage incident is old from 2016 in Peshawar

iVerify

Another trend observed by iVerify Pakistan was how differences between two sectarian groups were overplayed during particular news cycles, i.e. the 2024 violence in KP's Kurram and Parachinar. The result: heightened feelings of resentment and isolation. During this period, old visuals were frequently shared without context.

55 https://www.iverifypakistan.com/news/1000238

Similar sentiments were also witnessed towards Afghan refugees in Pakistan, particularly after the government's drive to repatriate them. That combined with the spike in terror incidents across the country — most blamed on Afghan nationals — the group became vulnerable to a hard online discourse. iVerify logged a total of seven leads with three pursued for further investigation that featured content against or highlighting the plight of Afghan refugees.
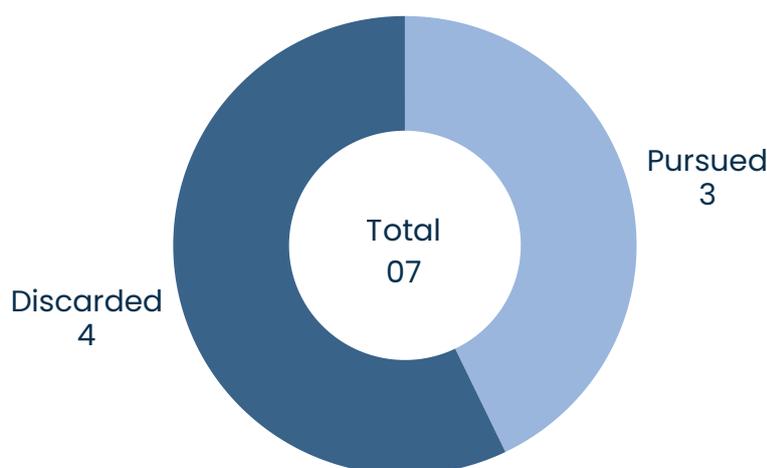
In April 2025, PML-N supporters[56] shared a video that allegedly showed the Sindh Police rounding up Afghan refugees with the help of ordinary citizens. The clip was celebrated as a positive occurrence.

On the other hand, those opposed to the repatriation drive shared content that showed visuals of alleged mistreatment[57] and excesses[58] against Afghan refugees in an attempt to highlight their plight and criticise the government. Regardless, content both for and against the repatriation had a harmful impact because they exaggerated the situation and possessed the potential to spread panic among the affected populace.

SOCIETAL 'VIRAL' INCIDENTS

iVerify saw how mis- and disinformation was created by sharp turns in public sentiment, stemming from intense and widespread emotions, especially anger over what members of the society perceived to be elite privilege, abuse of power, negligence of civic responsibilities, lack of accountability and other such societal topics. With such incidents involving private individuals, collecting information for verification can become almost near impossible due to the sensitivity of their personal details and security.

[56] https://www.iverifypakistan.com/news/1000289/video-shows-sindh-police-and-ancc-rounding-up-drug-users-not-afghans-2

[57] https://www.iverifypakistan.com/news/1000288

[58] https://www.iverifypakistan.com/news/1000476/viral-video-does-not-show-policemen-beating-up-afghan-women-2

SECTION 3

The Karsaz road accident served as a prime example of this trend. In August 2024, a speeding SUV, driven by a woman named Natasha Danish, killed a father-daughter duo on a bike who were identified as a sixty-year-old Imran Arif and 22-year-old Amna. The accident sparked furore over class differences after Natasha was identified to be the wife of a renowned industrialist.
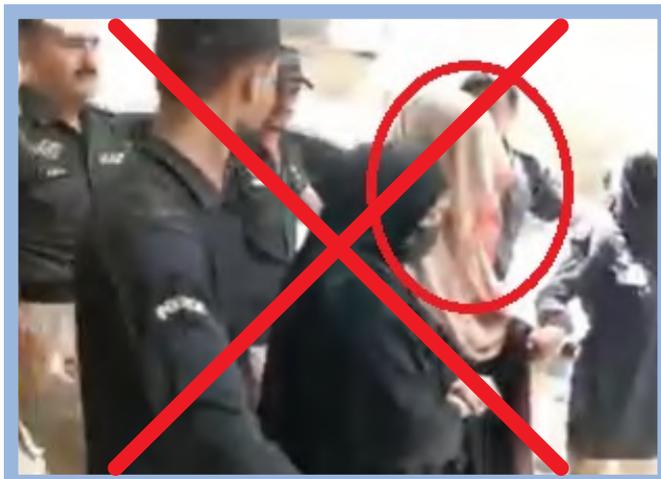
Throughout the case proceedings, social media was flooded with demands for justice and harsh criticism of the perpetrator. Eventually, the chatter reached a point where the debate turned into a case of 'poor vs rich', leading to rising unproven claims which couldn't be investigated conclusively as true or false. The media's role in this matter didn't help, with some publications altering or taking their news stories down altogether, adding to the anger.

Similarly, in October 2024, the news of a student's alleged rape at the Punjab Group of Colleges, Lahore, sparked an uproar on social media that came to the fore after students took to the streets to demand justice for the alleged victim. The news of the incident initially spread via digital news pages that reported on student testimonies. Later, the news took over mainstream news channels as the situation escalated following a clash between the protesting students and law enforcement officers.

Screenshots of messages narrating the incident of a guard and van driver allegedly raping a student on campus went viral and sparked a debate over students' safety in schools and how educational institutions covered up such incidents to protect their reputation.

The disarray wasn't helped when the PTI and PML-N used the situation to point fingers at each other. While PTI accounts and activists used the situation to highlight the state's negligence, Punjab Information Minister Azma Bokhari and Punjab CM Maryam Nawaz blamed PTI for spreading misinformation

and creating unrest in the country during the Shanghai Cooperation Organisation's (SCO) summit in Islamabad.

These examples show how public sentiment, the news cycle, and mis- and disinformation feed off and impact each other, ultimately creating a web of information so tangled that it poses challenges for fact-checkers to set the record straight.

3.2 COORDINATED CAMPAIGNS AND THE 'ANTI' BRIGADE

Coordinated campaigns are joint efforts by groups of perpetrators who share material from multiple accounts in large numbers, usually across multiple social media platforms, to push a narrative. Among the many topics observed and unpacked by iVerify, the following three demonstrate in particular how such campaigns operate. While mis- and disinformation cannot be quantified, the numbers presented in these campaigns show how coordinated amplification can distort public perception, making narratives appear credible, widespread, and widely endorsed.

BYC AND BALOCHISTAN

The issues plaguing Balochistan — chiefly human rights and missing persons — are among the most censored in Pakistan, especially because of the rise of militancy in the province. The clampdown on indigenous rights movements, such as Dr Mahrang's Baloch Yakjehti Committee, deepens the province's disconnect with other parts of the country.

The rise in militancy in Balochistan over the past few years has put organisations such as the BYC and the people associated with it at the centre of the vicious cycle of mis- and disinformation. An analysis of iVerify's pursued stories revealed that this spread of misinformation is a calculated effort to undermine the cause of the Baloch rights movement and link it with terrorism or as "anti-state".

SECTION 3

iVerify logged a total of 18 leads related to sentiment against (15) or in favour (3) of the Baloch community from December 2023 to November 2025.

One prominent example is the disinformation campaign falsely linking a Baloch woman, who was an aide of Dr Mahrang, to a violent incident involving a suicide bombing. Several social media posts circulated[59] in August 2024, claiming that she was the same person who carried out a suicide attack claimed by the outlawed Balochistan Liberation Army (BLA) in Lasbela. The incident drew attention due to the resurgence of militant activity in the region, and this false association was used as a tactic to equate Baloch activism with terrorism.

The misinformation campaign began with viral posts by prominent social media influencers and political figures who shared misleading content. One of the earliest claims came from Farhan Virk, a former PTI propagandist turned social media activist, who in a post on X alleged that the bomber involved in the Lasbela attack was the same woman who had been arrested by the Counter Terrorism Department (CTD) in February 2023 and subsequently released in May following public outcry. His post amassed over 126,000 views and 1,400 reposts.

Notably, content creator Saad Kaiser also made similar assertions in response to human rights lawyer Imaan Mazari-Hazir's posts, questioning her stance on the 2023 detention of the woman. Kaiser's post, which featured a side-by-side comparison of the alleged suicide bomber and the woman in question, garnered over 74,000 views. Other users echoed similar allegations, contributing to a significant online reach that further amplified the misleading narrative.

However, the Balochistan Counter Terrorism Department deputy inspector general of police and a visual analysis by iVerify Pakistan confirmed the two women to be different individuals.

---

[59] https://iverifypakistan.com/female-bla-suicide-bomber-of-lasbela-attack-is-not-the-same-mahal-baloch-arrested-by-ctd-in-feb-2023/

The mis- and disinformation in these instances primarily served two purposes: to deepen ethnic divisions between the Baloch community and the federal government, and to portray social rights movements as anti-state threats. The use of such harmful content is a significant tool in controlling the narrative in conflict-prone areas such as Balochistan, often at the cost of factual accuracy.

Dr Mahrang herself has been a victim of such coordinated campaigns, particularly during or in the aftermath of sensitive developments related to Balochistan.

A day after Jaffar Express hijacking on March 11, 2025 — when terrorists affiliated with the Balochistan Liberation Army ambushed and attacked the Peshawar-bound train carrying 440 passengers — PML-N supporters and PTI critics on social media shared a video of PTI General Secretary Salman Akram Raja, claiming[60] that he allegedly expressed support for BLA and Dr Mahrang. The aim was to conflate the terrorist outfit with the activist. In reality, Raja had not mentioned BLA at all and instead said the PTI would want to invite Dr Mahrang to Islamabad.

The same month, a picture of Dr Mahrang at her father's grave was also spread on social media with the claim that it was draped with the BLA's flag. This tied into a long-running claim observed on social media that made her father, Abdul Gaffar Langove, a BLA terrorist. The same picture was later shared in September 2025 by pro-military accounts across X, Facebook and Instagram with the same caption as part of a larger graphic that called her a terrorist's daughter.

The danger and harm of such campaigns and the snowball effect they can have in forming certain narratives can be judged by the military spokesperson himself labelling Dr Mahrang a "proxy" in May 2025 for terrorists. The labelling was criticised by veteran PPP politician Farhatullah Babar.

The most significant example in her case, that shows the methods such coordinated campaigns employ and how they operate, again came in September 2025 when multiple posts by pro-military accounts on social media cast aspersions on Jørgen Watne Frydnes, chair of the Norwegian Nobel Committee, by linking him to Zionism.

The posts alleged a nefarious connection between him and Dr Mahrang's nomination for the Nobel Peace Prize. They also used a screenshot of his Wikipedia profile that described him as a "staunch Zionist Jew". However, checking his Wikipedia profile showed no such description. Checking the edit history showed that a change[61] was made to his profile on September 24, 2025, at 7:58pm by an IP address that was traced to Islamabad. The coordinated posts against her were posted a few hours later, beginning from 12:49am on Sept 25.

IHC JUDGES AND THE LETTER

In the last two years, the judiciary too became a target of sustained and coordinated campaigns, as observed by iVerify. However, a more prominent coordinated campaign was seen against judges of the Islamabad High Court (IHC).

---

[61] https://en.wikipedia.org/w/index.php?title=J%C3%B8rgen_Watne_Frydnes&oldid=1313185637
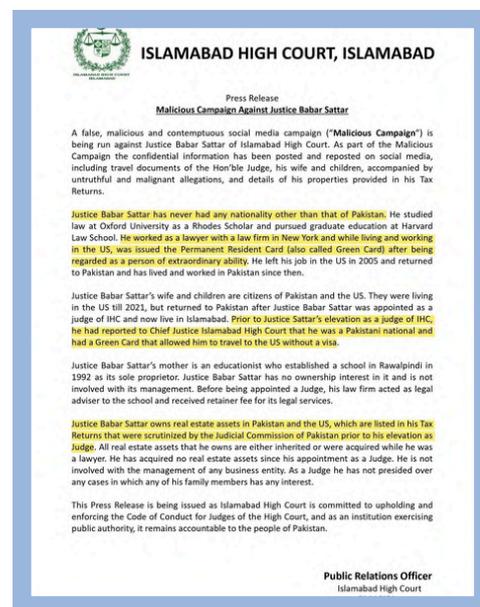
SECTION 3

On March 25, 2024, six judges of the IHC sent a letter to the Supreme Judicial Council (SJC), alleging interference from intelligence agencies. The judges who signed the letter were:

- Justice Mohsin Akhtar Kayan
- Justice Tariq Mehmood Jahangiri
- Justice Babar Sattar
- Justice Sardar Ejaz Ishaq Khan
- Justice Arbab Muhammad Tahir
- Justice Saman Rafat Imtiaz

This unprecedented move sparked a wave of misinformation, which was widely shared by various accounts on social media platforms. The PTI and PML-N in particular seized on the situation, each exploiting it to advance their own narratives. PTI supporters on social media claimed that the interference was evidence of state activity against the party. Meanwhile, PML-N supporters framed it as an example of the judiciary being compromised under pressure.

Following the letter's release, various claims circulated on social media, often lacking credible evidence and fuelled by political agendas, targeting the aforementioned judges.

On April 22, 2024, alleged travel documents of Justice Sattar surfaced online, accompanied by a claim that he had concealed[62] his US citizenship. The post garnered over 212,000 views, with the IHC swiftly denying it. The post originated from an X account, namely the Pakistan Lawyers Forum, which was later deleted, indicating a lack of accountability for those disseminating fake news.

Subsequently, on June 12, 2024, further misinformation emerged related to Justice Sattar when posts from pro-military accounts and PML-N supporters circulated purporting to show a statement from banned militant group Tehreek-i-Taliban Pakistan (TTP) praising him for a verdict regarding telecom companies and state surveillance the same year.

However, the statement was confirmed to be fabricated[63]. This claim was initially propagated by what seemed to be a pro-military account, gaining over 23,000 views, and was subsequently shared by journalist Gharida Farooqi, who is seen as supportive of the PML-N government and military establishment.

Justice Jahangiri was the next target. On July 5, 2024, a notification was circulated by a PML-N supporter on X claiming that the University of Karachi had declared the judge's LLB degree as fake. The alleged notification was viewed by more than 5,000 users. The team also observed that the account was followed by PML-N Senior Vice President and Punjab Chief Minister Maryam Nawaz.

iVerify reached out to the University of Karachi's administration for a comment, but despite multiple attempts, the varsity did not cooperate and refused to provide an official statement on the matter. Simultaneously, the claim continued to go viral and attracted widespread public interest.

[63] https://www.iverifypakistan.com/news/1000104

On July 9, 2024, journalist Gharida Farooqi once again shared allegations about Justice Jahangiri's degree being fake, claiming it had been confirmed by the registrar of the University of Karachi. Her post was viewed by more than 45,000 X users. In her post, the journalist did not provide a named source to support her assertions, relying instead on the ambiguous term "reportedly" in her statements. This led to the IHC issuing notices to Farooqi and another journalist for defaming the judge.

The matter culminated[64] in December 2025 when the judge was fired over the matter. It is thus seen how coordinated campaigns, particularly by groups aligned with the government, can act as a precursor to official action by first forming and channelling a certain narrative that subsequently lends legitimacy to any official action.

IMRAN AND INCARCERATION

The PTI's founder, due to his popularity and continued incarceration, was observed to be at the centre of many coordinated campaigns both by his party supporters and by those from the opposition camps.

An example of the campaigns run in his favour involved PTI supporters capitalising on the US elections to push for the former prime minister's release. In December 2024, a misleading post involving Richard Grenell, a US official and PTI sympathiser, gained significant traction on social media. Grenell had previously called for Imran's release, making him a figure of interest in Pakistani political discussions.

On January 9, 2025, an image purportedly showing Grenell calling for the resignation of army chief General Asim Munir circulated widely among Pakistani users on X. The post was quickly amplified by several PTI supporters, who interpreted the message as a direct demand for the army chief's resignation.

However, the image was a manipulation[65] of Grenell's original post, which referred to Los Angeles Mayor Karen Bass. Grenell had commented, "Resign, Karen", in response to a fire-related incident in Los Angeles.

In another case, PTI supporters shared a viral clip across social media platforms featuring Grenell praising Imran and calling for his release on January 3, 2025. However, the clip was dubbed over[66] as Grenell's original speech at the 2024 Republican National Convention focused on US President-elect Donald Trump.

Another example of a coordinated campaign regarding Imran, this one originating from foreign perpetrators, occurred in November 2025. Imran's relatives and lawyers were not allowed to meet him in prison for over a month, leading to concerns about his health.

---

65 https://www.iverifypakistan.com/news/1000251
66 https://www.iverifypakistan.com/news/1000248

On November 26, an Afghan propaganda account on X claimed that "a credible source from Pakistan" had confirmed to *Afghanistan Times* that PTI founder Imran Khan was mysteriously killed and his body moved out of prison.The post[67] gained 3.3m views and the claim quickly gained traction with it being reshared by several Indian and Afghan propaganda accounts.

Indian media outlets also covered the rumours, with Republic TV anchor Arnab Goswami amplifying the narrative in a show using the hashtag #WhereIsImranKhan. Other international publications, such as *Reuters*[68] and *Nikkei*[69,] also ran reports on the discussion.

PTI leadership and supporters, meanwhile, used this confusion to gain access to the leader — a matter of contention between the party and government — by releasing posts that demanded answers about his health, instead of offering clarity. The party in a post from its official X account put the onus on the government to clarify the matter while simultaneously giving statements to the media that there was no cause for concern.

The buzz over the status of Imran's health was thus seen to be strategically and calculatedly used by the PTI to increase pressure on the government to allow access to the incarcerated leader instead of deflating it by issuing a categorical party statement. The development eventually led to a meeting between Imran and his sister Uzma Khanum on December 2, 2025.

3.3 HATE SPEECH: DEEPENING ETHNIC/RELIGIOUS DIVIDE

While harmful content is broadly damaging for those it targets, content designated as hate speech goes a step ahead in the vitriol it expresses and the danger it holds.

The United Nations describes[70] hate speech as "any kind of communication in speech, writing or behaviour, that attacks or uses pejorative or discriminatory language with reference to a person or a group on the basis of who they are, in other words, based on their religion, ethnicity, nationality, race, colour, descent, gender or other identity factor".

[67] https://x.com/PTIofficial/status/1993782259946545183?s=20
[68] https://www.reuters.com/world/asia-pacific/imran-khan-healthy-cut-off-jail-sister-says-family-seeks-access-2025-12-02/
[69] https://asia.nikkei.com/politics/sons-of-pakistan-s-jailed-imran-khan-voice-fears-for-his-safety
[70] https://www.un.org/en/hate-speech/understanding-hate-speech/what-is-hate-speech

SECTION 3

A core function of iVerify since its launch has been to monitor such content due to its extreme nature; the team observed that harmful content and hate speech directed at political parties and figures often took on a religious tone as those spreading the incendiary material capitalised on the religiously charged sentiments of society.

For instance, posts perpetuating links between the Ahmadi community — referred to by the derogatory term 'Qadiani' — and PTI leaders were fact-checked and observed to spread only for the purpose of inciting hate. Hate speech was a mainstay in claims pertaining to the judiciary, particularly during times of politically charged cases in 2024. The team observed sustained campaigns against the courts, particularly against former chief justice of Pakistan Qazi Faez Isa or those judges considered to be biased against Khan by PTI supporters.

This campaign and environment of negativity towards the Supreme Court came to a head in February 2024 when JUI-F chief Fazlur Rehman[71] incorrectly accused Justice Isa of "allowing the Ahmadi community to spread its beliefs" after a court verdict.
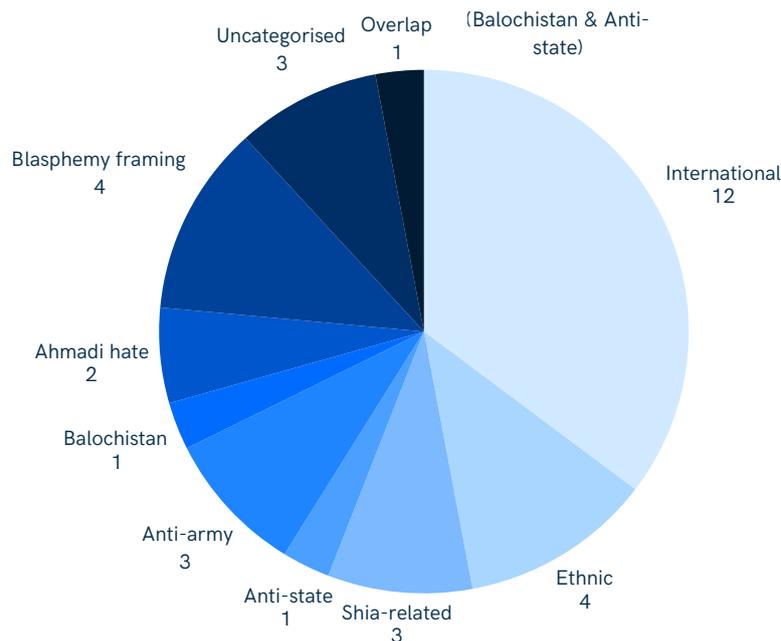
From 513 published stories, 34 were classified as hate speech after examining multiple angles; including local and international.

Twelve reports involved an international angle while four each concerned hate speech targeted on ethnic and religious grounds, three each were related to anti-military content and the Shia community while two each were related to targeting the Ahmadi community, and Baloch issues and anti-state labelling, with one story overlapping in the last two categories.

71 https://www.iverifypakistan.com/news/1000057

SECTION 3

In January 2025, posts from multiple users shared a video[72], claiming it showed supporters of Shia religiopolitical Majlis Wahdat-i-Muslimeen (MWM), an ally of the PTI, chanting anti-state slogans at a Karachi sit-in. In reality, the chants were actually against the United States, not Pakistan. The clip, posted by a pro-PMLN account, gained

76,500 views. The allegations were a clear fabrication to spread and incite hate against the religio-political group and possibly invite action from the authorities against it.

This tactic of vilifying individuals or institutions based on religious and/or political differences undermines democratic principles and pluralism, and most dangerously, escalates heightened sensitivities as in the case of Shia and Sunni communities.

Sectarian tensions between Shia and Sunni communities in Pakistan have long contributed to societal divisions, and these deep-seated grievances are frequently exploited through the spread of mis- and disinformation. The circulation of viral content, particularly around times of unrest, leverages historical grievances and fears, inflaming religious animosity.

Disinformation, whether through the recycling of old imagery or the misattribution of foreign events, thrives in the context of sectarian conflicts, leveraging pre-existing religious sensitivities and communal identities to stoke emotions and deepen divisions. By tapping into collective memory, it invokes the pain of past events to reignite contemporary tensions.

The iVerify team published seven reports that addressed harmful content that touched upon or was based upon sectarian issues.

---

[72] https://www.iverifypakistan.com/news/1000246

SECTION 3

On July 28, 2024, an image[73] depicting a large number of coffins was shared widely on social media, purportedly showing the aftermath of sectarian violence in Parachinar, a Shia-majority region in the Kurram district of Khyber Pakhtunkhwa. The area had witnessed deadly clashes that month, with 49 lives lost and over 200 people injured, raising alarm within the Shia community.

The image gained significant traction, receiving 19,500 views, with users expressing outrage over the alleged killings. In reality, the image was from a 2009 Karachi mock procession, but its circulation at a time of heightened tension in Parachinar exploited collective memory, fuelling fears of sectarian hate and violence. The re-use of old imagery was designed to evoke an emotional response and to create the perception of continuous violence in Shia audiences.

The widespread sharing of this image was not coincidental. It corresponded with unrest in Kurram which made the public more susceptible to misinformation. By evoking powerful imagery from a past event, the recycled content was strategically used to trigger a visceral response. The timing of the content's spread was crucial for maximising its impact, with the false narrative being amplified by users across various platforms to provoke outrage.

[73] https://iverifypakistan.com/viral-image-of-coffins-of-shia-victims-from-parachinar-violence-is-from-a-2009-karachi-mock-procession/

SECTION 3

The instability in Kurram in particular proved to be a sustained source of harmful content and hate speech that capitalised on sectarian tensions. As the crisis unfolded in the region, a surge of misinformation on social media exploited public emotions, fueling sectarian divides and complicating the situation. Below are notable examples:

On November 21, the same year, posts on social media platforms shared visuals of an infant, falsely claimed to be a Shia victim of the Kurram violence. Viewed by over 20,000 people, the image was fact-checked[74] and traced back to an older date with no connection to the recent violence.

On November 22, a video posted on X showed a man holding another person underwater in a shallow stream. The post alleged it was a Shia victim of the ongoing Shia-Sunni violence in Kurram. The video quickly went viral, garnering over 180,000 views.

However, investigation revealed[75] that the clip had been circulating for three years, often linked to unrelated conflicts in Afghanistan. Despite extensive efforts, no credible information confirmed its origin, making it clear that the video predated the violence in Kurram.

---

74 https://www.iverifypakistan.com/news/1000223
75 https://www.iverifypakistan.com/news/1000222

SECTION 3

The video's repurposing to mislead audiences about the situation in Kurram and inflame sectarian tensions highlights the harmful role of misinformation during such crises between various social groups.

Conversely, on November 24, a graphic image of a man holding a dead infant went viral, with posts alleging the child was killed, beheaded, and burnt in Bagan town by a Shia group. Our investigation revealed[76] the image dated back to 2016 and was from Afghanistan. It bore no connection to Kurram's situation at the time.


The deliberate use of fabricated content to fuel sectarian grievances highlights the dangerous role of mis- and disinformation in exploiting existing divisions and escalating tensions. Such narratives weaponise graphic and emotive content to provoke outrage, deepen mistrust between communities, and divert attention from addressing the root causes of violence. It further undermines peace efforts, posing a significant threat to stability in an already volatile region.

In these instances, the perpetrators of mis- and disinformation usually rely on recycled content and the strategic timing of its release to maximise impact. Misinformation that touches on sectarian issues is particularly effective because it taps into pre-existing religious identities and grievances. In the absence of real-time verification, these false narratives spread rapidly, with potentially dangerous consequences for inter-community relations.

The timing of these false stories is crucial. Whether it's during rights protests or sectarian clashes, misinformation is strategically inserted into the public discourse at moments when people are most likely to react emotionally and share content without questioning its authenticity. This ensures maximum impact, making it harder for fact-checkers to correct the narrative once it's taken hold.
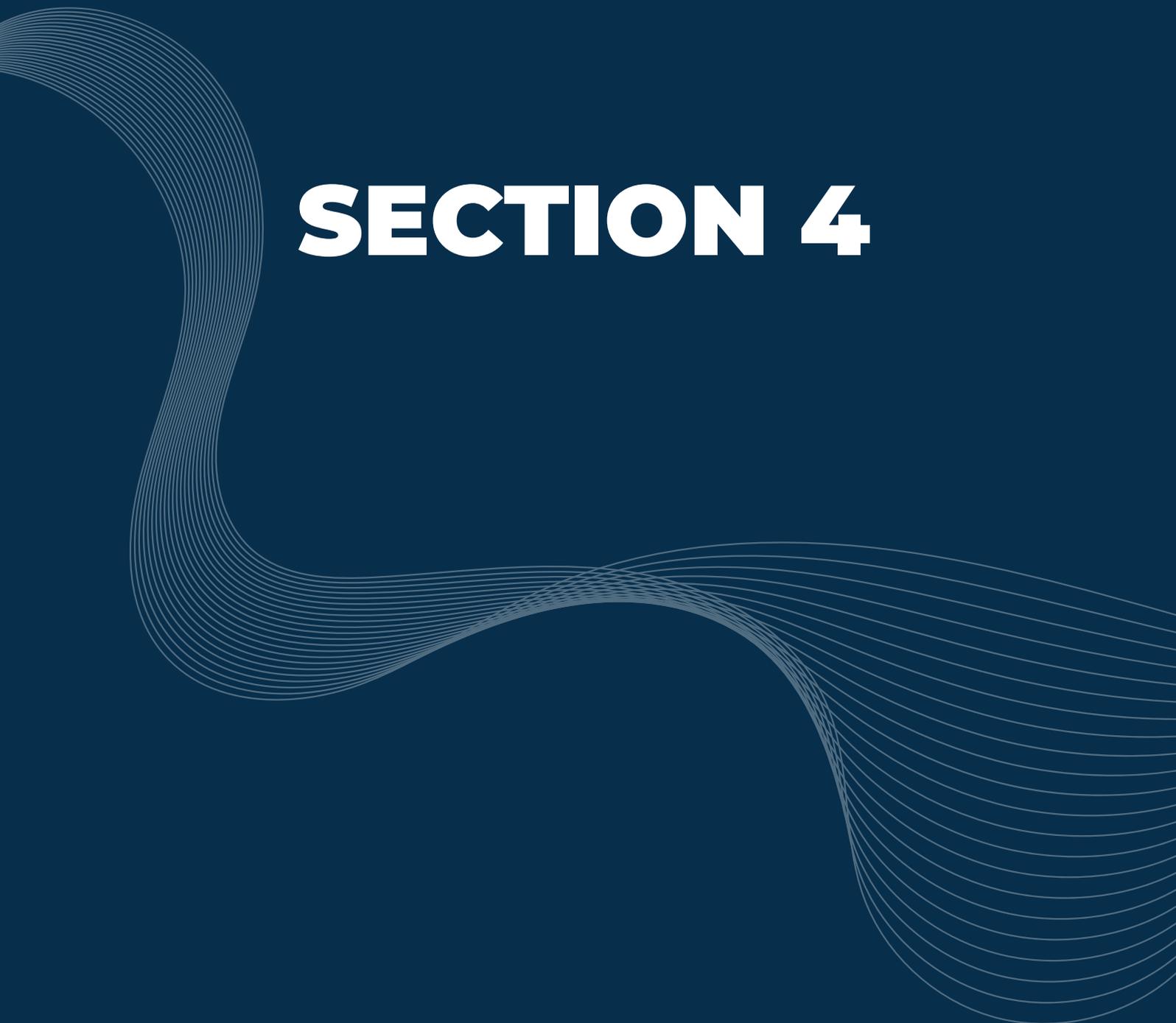
76 https://www.iverifypakistan.com/news/1000224

SECTION 4

CHALLENGES

Fact-checking as an enterprise faces many of the same challenges that the wider media landscape faces in Pakistan: lack of resources, lack of cooperation from state officials and institutions on certain topics, especially those related to security and the military, and lack of official statements by parties involved that can aid in completing a story. Similarly, investigative efforts are adversely affected by internet and social media clampdowns as witnessed in 2024, during and after elections.

Based on iVerify's two-year work on evidence-based fact-checking, the main challenges to countering mis- and disinformation are listed below:

VOLUME AND SPEED

The biggest challenge in fact-checking remains timing and delivery: spreading a credulously fabricated visual with an accompanying caption can take a mere click of a button while the process of investigating it is a longer exercise.

By the time a fact-check, with clearly laid out evidence, is published, the piece of false information would have spread to various platforms and already achieved the desired effect. This matter is particularly difficult to address during periods of heightened activity – such as elections, political unrest or war/conflict – when the sheer influx of information on social media platforms makes it difficult to track all claims.

Thousands of users sharing different videos and visuals with varying claims necessitates greater editorial decision-making and application of criteria to pick and choose the leads requiring further investigation.

RECIRCULATION OF FACT-CHECKED CLAIMS

The recurrence of mis- and disinformation, whereby the publication of a fact-check report does not itself put an end to the spread of a piece of content, also remains a major issue. It has been observed that certain perpetrators spread the same old images and videos with the same claims repeatedly, mostly over a period of time, while accruing high views. Fact-checking a piece of content, therefore, not only requires investigating it once but also regular monitoring to preemptively respond to it when it is shared next.

ECHO CHAMBERS

Closed networks of like-minded individuals reinforcing each other's beliefs. Among such groups, when a piece of mis- and disinformation is introduced, it spreads rapidly without scrutiny, as the members of the echo chamber tend to share content that aligns with their existing views, regardless of its accuracy. These echo chambers continue to push disinformation, even after being debunked, creating cycles of fallacious content spreading. If public figures are commandeering these echo chambers, the information becomes even harder to dissuage.

SPOTTING AI

Another major challenge that dominates the focus of fact-checking efforts is related to the investigation and analysis of AI-generated content. Earlier attempts at such visuals had certain telltale signs or errors that made it easier to spot them and point them out to audiences. However, with the march of technology and increasing sophistication of newer programmes, AI-generated content is fast approaching a level of detail and quality that overcomes many of these earlier errors.

71

LACK OF TOOLS TO TACKLE AI

Compounding the aforementioned issue is the lagging reliability of tools to flag and detect AI-generated images and videos, with no single multipurpose tool available yet that provides a high degree of reliability for all cases. Investigating through the use of forensic tools thus often requires testing on multiple tools, with the primary method being to still look for visible flaws in a piece of content or attempt to dismantle its logic. This level of detailed and complex investigation adds to the issue of time as the fact-check tackling a claim of such nature would take even longer.

OVERRELIANCE ON AI AND VERIFICATION TOOLS SUCH AS GROK

Listing social media as a challenge would be too broad but many challenges link back to its usage which has made users comfortable with quick methods for consuming news and now, verification of news. As the Executive Summary notes, the lack of faith in media adds to this factor but the future could decide if people revert to mainstream media for verified information. In the meantime, however, people are turning to free AI models that can offer incomplete or biased information[77], and are also prone to hallucinating. AI bots such as Grok that are flawed[78] in their methodology, results and use are integrated into the social media platform, making it the most accessible verification platform for X users.

[77] https://lib.guides.umd.edu/c.php?g=1340355&p=9880574
[78] https://www.theguardian.com/technology/2025/nov/03/grokipedia-academics-assess-elon-musk-ai-powered-encyclopedia
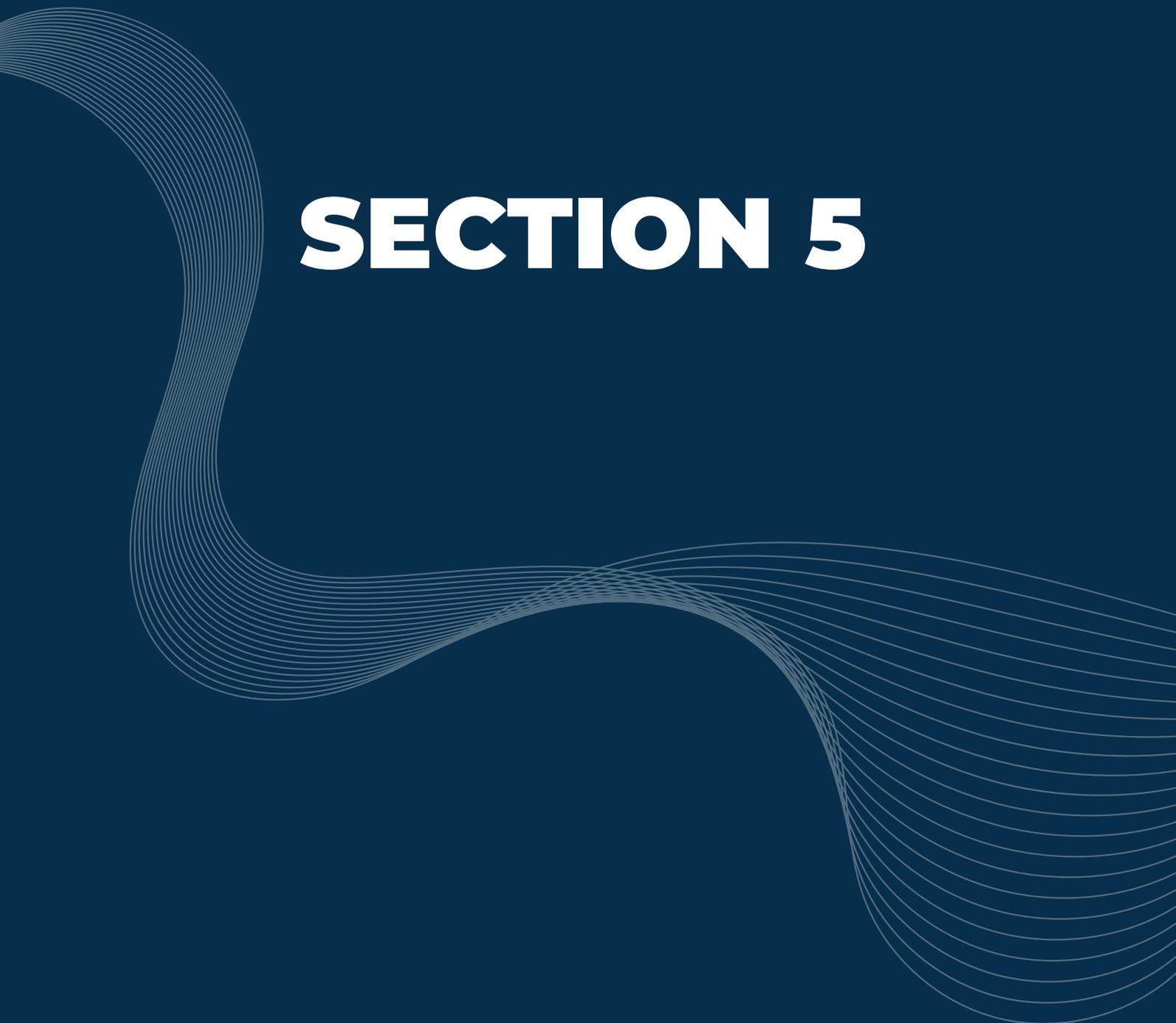
CLAIMS FROM STATE AUTHORITIES

As exemplified during the India-Pakistan military actions in May 2025, fact-checking faces challenges in times of conflict with no access to sites of combat, wide-ranging statements from authorities on both sides which can not be verified independently and officials quoting[79] false information that further perpetuates it.

PRESENCE OF MULTIPLE VERIFIED ACCOUNTS

The presence of numerous verified (blue tick) accounts on X creates confusion even among fact-checkers as the subject cannot always be contacted to confirm a matter. This also poses a challenge for fact-checkers when quoting official statements as evidence, requiring an extra layer of verification.

79 https://www.boomlive.in/fact-check/pakistan-deputy-prime-foreign-minister-ishaq-dar-telegraph-report-undisputed-king-air-force-fact-check-28589

SECTION 5

LOOKING AHEAD: HOW FACTS CAN SURVIVE MIS- AND DISINFORMATION AND AI

Many of the challenges above circle back to a lack of resources in fact-checking, technical complexities and the role of media as mentioned in the Executive Summary under Section 1. To resolve these, collaboration among stakeholders can play a pivotal role.
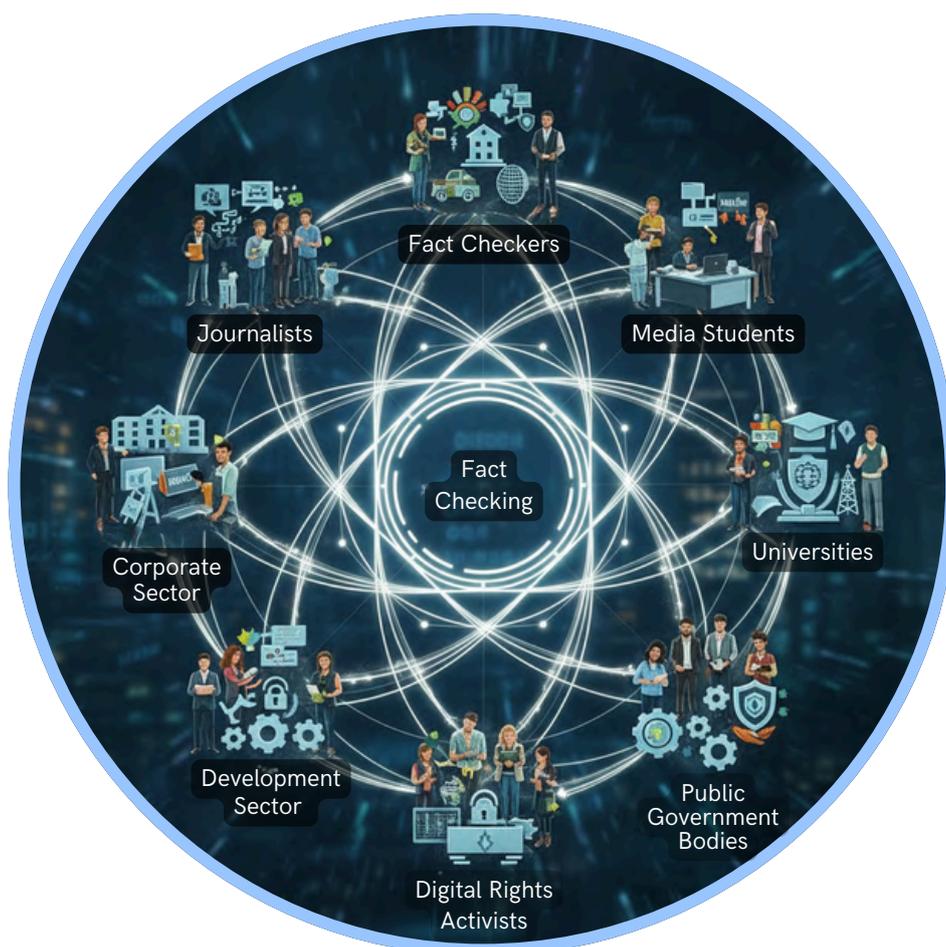
Countering misinformation in all its forms, across all mediums is a responsibility that has to be shouldered by all stakeholders.

SECTION 5

Below are recommendations grounded in iVerify's findings and experiences from December 2023 to November 2025 based on the patterns, risks, and challenges identified across the report's sections.

GOVERNMENT BODIES

- Upgrade public communication mechanisms, prioritising bilingual (Urdu–English) social media campaigns ahead of elections; explain voting procedures, address FAQs and debunk viral falsehoods in real time.

- Coordinate proactively with mainstream media and fact-checking organisations to ensure rapid verification during high-activity periods, such as climate emergencies and regional conflicts.

- Oppose blanket internet shutdowns and develop contingency communication plans in coordination with civil society if disruptions are imposed for security measures.

- Recognise AI-enabled misinformation as a high-risk category, requiring faster review, higher scrutiny and specialised response protocols.

- Prioritise protection against gendered disinformation, including sexualised deepfakes targeting women in politics, journalism and public life.

MEDIA/JOURNALISTS

- Avoid amplifying unverified claims that label groups as 'anti-state', especially during protests, rights movements or security operations.

- Institutionalise fact-checking in daily newsroom workflows, either through dedicated teams or partnerships with credible third-party initiatives such as iVerify.

- Discourage premature "breaking news" on personal social media accounts, particularly when information is incomplete or source-based.

- Exercise restraint in resharing viral content, even with disclaimers; journalists' amplification carries disproportionate influence.

- Limit the use of anonymous "sources", especially on social platforms, and issue prompt corrections when errors occur.

- Run sustained public awareness campaigns — in Urdu, English and regional languages — on how audiences can identify and question misinformation.

- Clearly label political advertisements, ensuring transparency and balance during election periods.

DIGITAL RIGHTS ACTIVISTS & CIVIL SOCIETY

- Strengthen collaboration between fact-checkers, human rights groups and conflict researchers to contextualise and counter divisive narratives.

- Advocate consistently against internet and social media shutdowns, highlighting their impact on voting rights, access to information and economic activity.

- Coordinate legal, media and public campaigns to document and challenge unlawful disruptions.

- Engage early with the ECP and government bodies to stress how uncertainty around connectivity undermines electoral integrity.

- Create public reporting mechanisms for documenting internet shutdowns and access issues.

SOCIAL MEDIA PLATFORMS

- Enforce stricter independent transparency and labelling of political advertisements.

- Work closely with local fact-checkers, election bodies and researchers, including fast-track review mechanisms during elections and crises.

- Proactively amplify verified fact-checks and official civic information, especially from trusted partners such as iVerify and the ECP.

- Open up collaboration with digital rights activists and international development bodies on policy consultation and devising mechanisms to check spread of mis- and disinformation.

CITIZENS

- Verify information through official sources and mainstream media before sharing.

- Be cautious of content citing unnamed "sources", especially on personal social media accounts.

- Look for missing context, half-truths or edited visuals, particularly during crises.

- Educate yourselves on responsible digital consumption and engagement of information. Here is a good start.

- Send claims to fact-checking initiatives like iVerify.

GLOSSARY

GLOSSARY

Agendas: Motives or goals pursued by individuals or groups.

Alleging/Allegations: Claiming or asserting something without proof.

Allegedly: Used to convey that something is claimed to be the case or to have taken place, although there is no proof.

Apostasy: The act of renouncing or abandoning one's religion or faith.

Blasphemy: The act of showing disrespect, insult, or irreverence toward God, religion, religious symbols, or sacred beliefs.

Bilingual campaigns: Campaigns conducted in two languages, typically English and a regional language, to ensure wider reach and accessibility, often used during election awareness campaigns.

Breaking news: News events that are happening in real time or have just occurred and require immediate reporting due to their significance or urgency.

Campaign: An organised effort to achieve a particular goal, especially in a political context, such as promoting a candidate for office or advocating for a specific policy.

Caretaker government: A temporary government appointed to administer a country during the period between an outgoing government and the incoming government after an election.

Claim: A statement or assertion made in text, audio, video, or image form that presents information as fact and can be investigated for accuracy.

Context: The circumstances or background surrounding a particular event or situation, providing a fuller understanding.

Constitution: A set of fundamental principles or established precedents according to which a state or organisation is governed.

Conviction: The act of finding somebody guilty of a crime in a court of law.

Debunked: Proven to be false after careful verification using evidence and facts.

Deepfake technology: The use of artificial intelligence, particularly deep learning techniques, to create or alter audio, images, or videos in a way that makes people appear to say or do things they never actually did, often with the intent to deceive.

Deepfakes: Digitally altered audio or video recordings that present a person doing or saying something they never actually did or said.

Deepfake videos: Synthetic or manipulated videos created using artificial intelligence to realistically alter a person's face, voice, or actions, making it appear as though they said or did something they never did.

Digital manipulation: The alteration of digital content, such as images, videos, audio, or text, using software or artificial intelligence to misrepresent, fabricate, or distort information in a misleading way.

Digital rights activists: Individuals or groups advocating for the protection of rights related to digital technologies and the internet, including freedom of speech and privacy.

Discourse: Communication or conversation involving a particular topic or subject. Disinformation: False or misleading information deliberately created and spread to deceive or manipulate.

Discarded claims: Statements or pieces of information that a fact-checking team decides not to investigate further because they are irrelevant, lack sufficient impact, or cannot be verified.

Dubbed videos: Videos in which the original audio is replaced with altered or new sound, making it appear that the person is saying something they never said.

Diyat: Commonly known as "blood money"; financial compensation payable to a victim or their next of kin in cases such as homicide or bodily harm, as an alternative to retribution (qisas).

Educational institutions: Schools, colleges, and universities that provide formal education and training.

Electoral body: An organisation responsible for administering and overseeing elections in a country, such as the Election Commission of Pakistan (ECP).

Election symbol: A symbol allotted by the election commission to a political party or independent candidate for use on ballot papers.

Electoral process: Procedures involved in conducting elections, including voter registration, campaigning, polling, and vote counting.

Electoral systems: Structures and mechanisms governing how elections are conducted, including voting methods and electoral laws.

Electorate: All people in a country or area who are entitled to vote in an election.

Ethical: In accordance with moral principles or standards.

Fact-checking: The process of verifying the accuracy and truthfulness of claims, statements, or information, particularly in news and public discourse.

Fake news: Information presented as news that is intentionally misleading or inaccurate and often spreads rapidly through media channels, especially during elections.

Freedom of speech: The right to express opinions and ideas freely without censorship or restraint; a fundamental democratic principle.

General elections: Nationwide elections in which eligible citizens vote to choose members of the legislature or public officials.

Imam: A religious leader or prayer leader in a mosque, particularly within Islam.

Impersonation: Pretending to be another person, typically for deceptive purposes.

Inaccurate information: Information that is incorrect or imprecise.

Inflamed religious sentiments: Heightened emotional reactions related to religious beliefs or practices.

Integrity: The quality of being honest, fair, and morally upright.

Internet shutdown: Government-imposed restrictions on internet access, often used to control information flow or suppress dissent.

Khatm-i-Nabuwat (ختم نبوت): The belief in the finality of Prophet Muhammad (PBUH).

Lead: A piece of information that shows potential for further investigation.

Lewd video: Content that depicts or implies sexually suggestive, indecent, or obscene behaviour intended to provoke sexual interest or shock, without necessarily showing explicit acts.

Manipulation: The deliberate distortion or alteration of information to mislead or deceive.

Media cell: A specialised unit, such as within the ECP, responsible for communication and media relations.

Media organisations: Entities that produce and disseminate news and information, including newspapers, television channels, and digital platforms.

Misinformation: Incorrect or misleading information spread without the intent to deceive.

Misinterpretation: An incorrect understanding or explanation of information. Narrative: A story or account of events shaped by a particular perspective or interpretation.

Nomination papers: Official documents submitted by individuals seeking to contest elections, including personal details and declarations.

Official authorities: Government bodies or agencies responsible for enforcing laws and regulations.

Patwari (پٹواری): A derogatory term commonly used for supporters or voters of the Pakistan Muslim League–Nawaz (PML-N).

Pervasiveness: The quality of being widespread or prevalent.

Political ads: Advertisements designed to promote a political candidate, party, or ideology, particularly during election campaigns.

Political parties: Organisations that seek political power and influence government policy through elections.

Prebunk: To prevent mis- and disinformation by educating or warning audiences before false content spreads.

Press freedom: The ability of journalists and media organisations to report freely without censorship or undue influence.

Propaganda: Biased or misleading information used to promote a political cause or ideology.

Provision: A specific clause or stipulation within a law or regulation.

Proxy: The authority to represent another person, especially in voting.

Public opinion: The collective views and attitudes of the general population on various issues.

Pursued claims: Statements or information that a fact-checking team has chosen to investigate further.

Qadiani (قادیانی): A religious slur used to refer to members of the Ahmadi community.

Religious orthodoxy: Adherence to established, traditional, and widely accepted doctrines or practices within a religion.

Sermons: Religious addresses or speeches delivered by clergy during worship or religious gatherings.

Sharia: Islamic law derived from the Quran and the traditions of the Prophet Muhammad.

Skirmishes: Minor or brief conflicts or engagements.

Social media clampdowns: Deliberate efforts by authorities to restrict or limit access to social media platforms, often during sensitive periods such as elections.

Stakeholder: Any individual, group, or organisation that has an interest in, is affected by, or can influence a particular decision or outcome.

Stonewalled: Refused to cooperate or provide information, often to delay or obstruct.

Synthetic media: Visual, audio, or textual content that is partially or fully generated, altered, or modified using artificial intelligence or machine learning.

Tickers: Scrolling text displayed on screens, typically providing news updates or key information.

Transcription: The process of converting spoken language or audio into written text.

Verification: The process of confirming the accuracy or truthfulness of information.

Virality: The tendency of content, especially on social media, to spread rapidly and widely.

Youthia: A derogatory term commonly used for supporters or voters of the Pakistan Tehreek-i-Insaf (PTI).