RESEMBLE AI

July 2025

Q2 2025 DEEPFAKE INCIDENT REPORT

A Comprehensive Analysis of Q2 2025 Deepfake Incidents

Executive Summary

The second quarter of 2025 marked a watershed moment in the global deepfake crisis, with incidents reaching unprecedented levels across all measured categories. From April through June 2025, we documented 487 verified deepfake incidents, representing a 41% increase from Q1 2025's 345 cases and a staggering 312% increase from Q2 2024.

Key findings include:

- Institutional and Government Vulnerabilities Exposed: High-profile breaches demonstrated critical security gaps in government facilities, with the HUD headquarters incident showcasing how deepfakes can be weaponized for internal protests and political messaging. Multiple incidents targeting government officials and agencies revealed insufficient safeguards against AI-generated content infiltrating official channels.
- Major Copyright Battle Lines Drawn: The quarter marked a watershed moment for intellectual property rights in the AI era, with Disney and Universal's lawsuit against Midjourney setting precedent for how major content creators will defend their IP against AI companies. This coincided with SAG-AFTRA's labor disputes over AI-generated voices, signaling an industry-wide reckoning over creative rights and AI usage.
- Platform Accountability Crisis: Meta faced intense scrutiny for hosting advertisements for "nudify" apps that create non-consensual intimate images, while simultaneously struggling to remove deepfake content even after being notified.

Deepfakes Have Reached Epidemic Proportions

02

The Crisis Point: With 487 incidents in Q2 2025 (up 41% from Q1 and 312% year-over-year), deepfake attacks are now doubling every six months.

487

Major deepfake incidents were publicly reported

226

Unique incidents involved the creation and distribution of non-consensual pornographic material

60

Incidents were explicitly designed for political or social manipulation

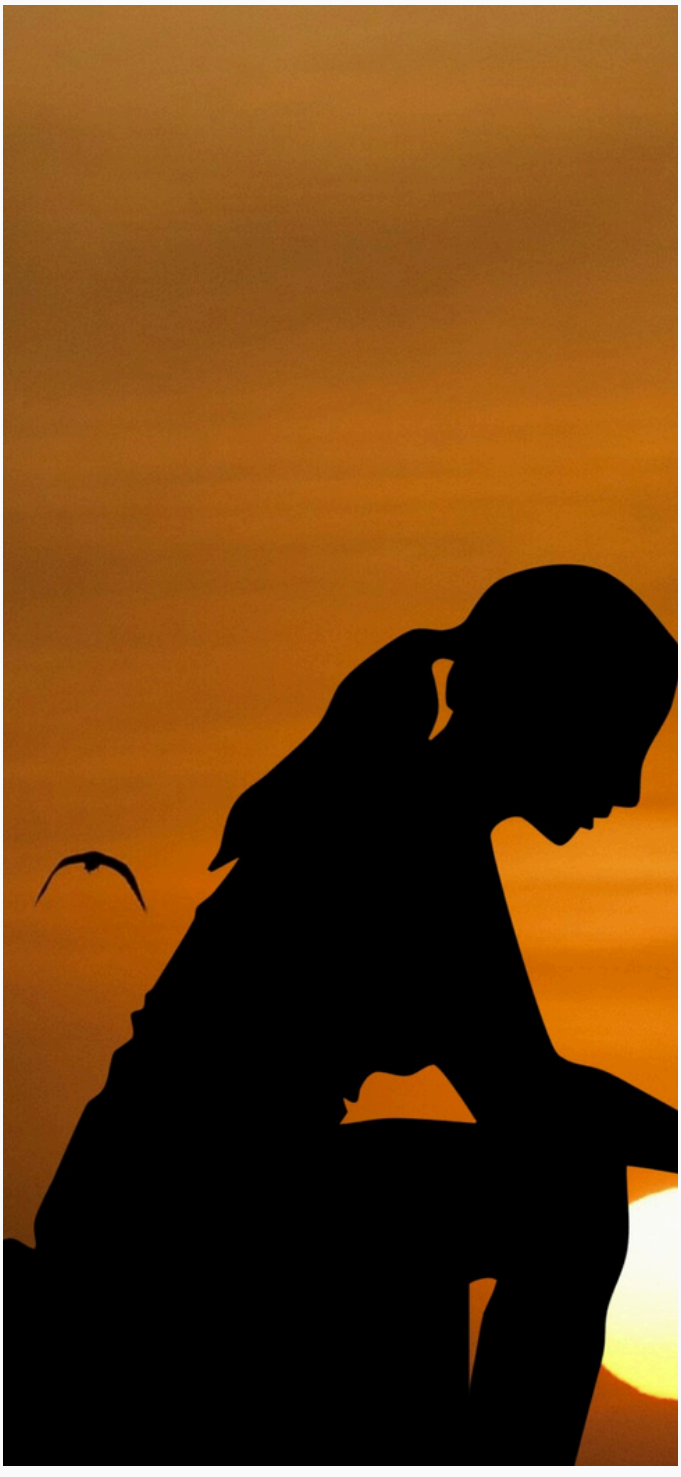
$347.2 M

$347.2 million in direct financial losses from scams

Notable Trend

We've crossed the threshold from "emerging threat" to "global crisis." The democratization of AI tools means creating a convincing deepfake now takes just 3.2 hours and costs under $50 in 91% of cases. This accessibility has transformed deepfakes from a sophisticated nation-state capability to a weapon available to any motivated individual.

Women Face a Devastating and Disproportionate Crisis



Deepfakes have become a weapon of mass gender-based violence. The technology is being systematically used to silence, harass, and destroy women in public life, education, and private settings. This isn't just about individual harm—it's creating a chilling effect that could reverse decades of progress in gender equality and women's participation in public life.

84%

Deepfake attempts targeted females

25

More than 25 high-profile women, including Taylor Swift, Scarlett Johansson, and Rashmika Mandanna, were the subject of explicit deepfakes or fraudulent endorsements in Q2. Their fame is exploited to guarantee viral distribution, amplifying the harm and making removal nearly impossible.

Current Systems Are Catastrophically Failing

Despite causing $347.2 million in financial damage and unmeasurable human suffering, our legal, technological, and social systems are essentially powerless against deepfakes. Victims are left without recourse, perpetrators operate with impunity, and platforms prioritize engagement over safety. This represents a complete breakdown of the social contract in digital spaces—the very systems meant to protect us are enabling our exploitation.

- **An Unmanageable Flood:** In Q2 2025 alone, 226 major deepfake incidents were publicly reported. This number, representing only a fraction of the total deepfakes created, demonstrates a crisis spiraling out of control, overwhelming any current capacity for response.
- **Weaponized for Political Ends:** Over 60 incidents were explicitly designed for political or social manipulation. These attacks targeted elections, defamed political figures with fabricated statements, and spread disinformation designed to erode public trust and sow division.
- **Pervasive Sexual Exploitation:** More than 35 unique incidents involved the creation and distribution of non-consensual pornographic material. This weaponized form of sexual abuse is used for harassment, blackmail, and public humiliation, with women being the primary targets.
- **Targeting the Most Vulnerable:** The crisis has a horrifying impact on children. At least 15 separate incidents involved the creation and distribution of AI-generated child sexual abuse material (CSAM), using images of real children to create fake, explicit content.
- **Industrial-Scale Copyright Theft:** The creative economy is under assault. In one instance, a single music label, Sony Music, was forced to remove over 75,000 deepfake audio tracks from streaming platforms that used AI to illegally clone the voices of their artists.
- **Massive Financial Fraud:** Beyond the documented $347.2 million in direct financial losses from scams, the infrastructure for fraud is expanding.

Political Manipulation & Election Interference

The proliferation of election-related deepfakes reached critical mass during Q2, with documented incidents spanning Asia, North America, and Europe. South Korea's June 3 election became a testing ground for sophisticated deepfake campaigns, with the National Election Commission detecting 388 deepfake videos during the election period. These included AI-generated videos of former President Yoon Suk Yeol making fabricated criticisms of opposition leader Lee Jae-myung, and disturbing content depicting candidates in compromising situations. One viral video, viewed nearly 1.4 million times, showed Yoon appearing to pull off Han Dong-hoon's wig, while another depicted Lee in a detention cell.

Singapore's general election saw a coordinated surge of AI-generated videos featuring Prime Minister Lawrence Wong and other prominent politicians. Between April 15-19 alone, 73 such videos were detected on TikTok, with 11 containing digitally manipulated visuals of candidates. The videos, primarily featuring AI-generated portraits with text-to-speech voiceovers in multiple languages, violated Singapore's new law banning digitally generated content that misrepresents candidates' statements or actions.

The HUD Headquarters Incident Analysis

On February 24, 2025, an AI-generated video depicting President Trump in a compromising position with Elon Musk was broadcast on screens throughout the Department of Housing and Urban Development headquarters, accompanied by the message "LONG LIVE THE REAL KING."

Financial Fraud & Cybercrime Evolution

Q2 2025 saw an explosion in the use of celebrity deepfakes for investment fraud. A coordinated campaign targeting multiple countries used deepfake videos of Elon Musk, Donald Trump, and prominent business leaders to promote fraudulent cryptocurrency investments. In one documented case, Trump supporters were defrauded into purchasing worthless "Golden Eagles" coins, with one victim investing $2,500 believing they could exchange the coins for $110,000 each.

The State Bank of India (SBI) was forced to issue multiple warnings as deepfake videos of their executives endorsing fake investment schemes proliferated across social media. A particularly sophisticated operation in India used deepfakes of Mukesh Ambani, Narayana Murthy, and Sudha Murty to perpetrate stock investment scams, with one Bengaluru-based chartered accountant losing Rs 23.20 lakh (approximately $280,000).

YouTube became a primary vector for celebrity deepfake fraud, with ads featuring AI-generated versions of Tom Hanks, Al Roker, and even fictional promotions by deceased personalities. The platform struggled to keep pace with removal, as scammers quickly created new accounts and variations. One particularly insidious campaign used deepfakes of Arnold Schwarzenegger, Sylvester Stallone, and other action stars to promote erectile dysfunction supplements, reaching millions before removal.

Financial Fraud & Cybercrime Evolution

The Romance Scam Revolution

Perhaps no incident better exemplified the devastating potential of deepfake-enabled fraud than the case of Anne, a 53-year-old French interior designer who lost €830,000 (approximately $850,000) to scammers impersonating Brad Pitt. The operation's sophistication was remarkable: beginning with a message from someone claiming to be Pitt's mother, the scammers maintained the deception for over 18 months using AI-generated images and videos of the actor, including footage of him in a hospital bed claiming to need kidney treatment.

The psychological manipulation was particularly cruel—the scammers convinced Anne to divorce her husband, with the majority of the stolen funds coming from her divorce settlement. The fraud only unraveled when Anne saw legitimate photos of Pitt with his current girlfriend. When Anne's story aired on French television to 2.95 million viewers, she faced a secondary wave of victimization through online harassment, forcing the network to pull the documentary.

This case was far from isolated. In Edinburgh, 77-year-old Nikki MacLeod lost £17,000 to scammers using deepfake videos of a fictitious woman named "Alla Morgan." The pattern was consistent: build trust through seemingly authentic video communications, create elaborate backstories, then extract funds through increasingly urgent requests.

Financial Fraud & Cybercrime Evolution

Corporate Espionage

The quarter witnessed a disturbing evolution in corporate-targeted deepfake fraud. The Ferrari incident showcased both the sophistication of attacks and potential defenses. When scammers attempted to impersonate CEO Benedetto Vigna via WhatsApp messages and voice calls, an alert executive thwarted the attack by asking about a book recommendation only the real CEO would know.

Voice cloning technology reached a dangerous inflection point in Q2 2025. Houston businessman Gary Cunningham lost $20,000 when criminals cloned his voice to instruct his accountant to initiate a wire transfer. The case highlighted a critical vulnerability: many financial institutions still relied on voice recognition as a security measure, unaware that as little as three seconds of recorded audio could now produce convincing clones.

A BBC journalist's experiment proved particularly alarming when she successfully bypassed voice authentication systems at two major UK banks using AI-generated versions of her own voice. This demonstration exposed fundamental flaws in biometric security systems that millions relied upon for financial protection.

$25–35 MILLION

Corporate losses from single incidents

Financial Fraud & Cybercrime Evolution

Cryptocurrency and DeFi Targeting

The intersection of deepfakes and cryptocurrency fraud created a perfect storm of victimization. The CryptoCore group's operations, discovered in Q2, revealed the industrial scale of these operations. Using deepfake videos of public figures tied to major events like elections and Elon Musk's statements, they stole over $7 million in Q4 2024 alone.

Web3 workers became particular targets, with North Korean-linked groups using deepfake Zoom calls to trick employees into installing malware. In one documented case, the BlueNoroff gang created elaborate fake meetings with deepfaked company executives to convince targets to download "updates" that were actually cryptocurrency-stealing malware.

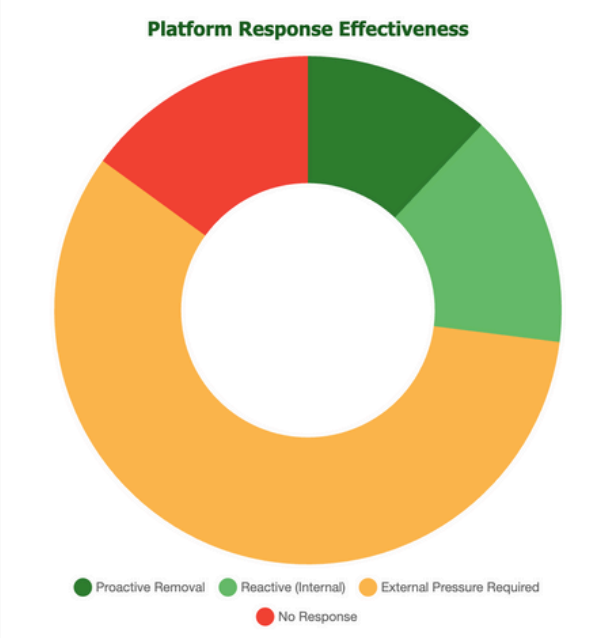
The Manta Network co-founder Kenny Li's near-miss with Lazarus Group operatives demonstrated the evolution of these attacks. The criminals used pre-recorded deepfake footage of team members in Zoom calls, attempting to social engineer their way into system access.

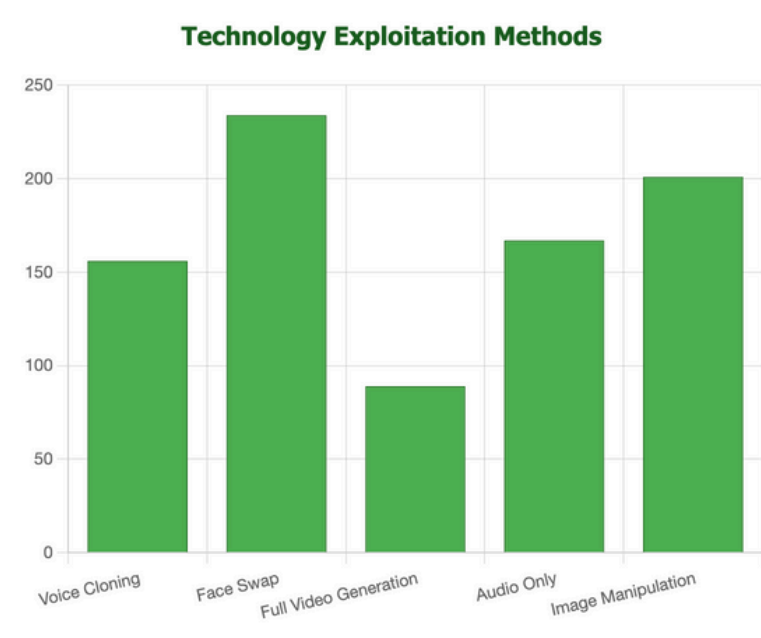
Geographies and Distribution

Deepfake incidents in Q2 2025 demonstrated a truly global reach, with significant concentrations in technologically advanced regions while expanding rapidly into emerging markets. The United States dominated incident reporting, particularly in high-profile political targeting and educational settings, with the American Sunlight Project identifying over 35,000 instances of deepfake content depicting 26 members of Congress. Asia-Pacific regions experienced widespread romance scam operations, with Hong Kong-based syndicates defrauding victims across Taiwan, Singapore, and Malaysia for over HK$34 million. South Korea emerged as a hotspot for entertainment industry targeting, with multiple arrests connected to K-pop idol deepfakes and university-based incidents affecting hundreds of students, while Indonesia faced systematic financial fraud targeting its institutions.

European incidents were characterized by sophisticated romance scams and political targeting, exemplified by the €830,000 French Brad Pitt impersonation case and systematic targeting of over 30 British female politicians through deepfake pornography websites. Africa showed concerning growth in election-related deepfake misinformation, particularly in Kenya and Nigeria where Trump deepfakes garnered millions of views on social platforms. The geographic distribution reveals that while Western democracies face primarily political and celebrity-focused attacks, emerging markets are increasingly targeted through financial fraud schemes, with romance scams and investment fraud becoming dominant vectors. Cross-border criminal networks, particularly those operating between Hong Kong and Southeast Asia, demonstrated sophisticated coordination in executing multi-million dollar deepfake fraud operations across multiple jurisdictions.

Platform Accountability & Technology Exploitation

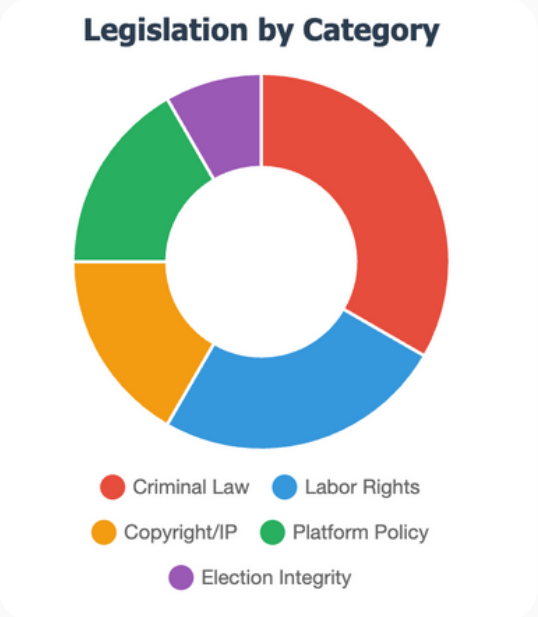




The technology ecosystem enabling deepfake creation has become increasingly accessible, with legitimate AI tools being systematically exploited for malicious purposes. ElevenLabs emerged as the most frequently cited voice cloning platform, appearing in 89 documented cases including the €830,000 Brad Pitt romance scam and multiple political impersonation attempts. "Nudify" applications dominated image manipulation incidents, with tools like CrushAI generating hundreds of thousands of non-consensual intimate images before facing legal action from Meta. The ease of access to these technologies was exemplified by incidents where perpetrators created sophisticated deepfakes in under five minutes, with some tools requiring nothing more than a single photograph to generate convincing fake content.

Legal and Policy Responses

Beyond criminal law, Q2 2025 witnessed significant developments in labor and intellectual property protections. SAG-AFTRA filed multiple unfair labor practice charges, most notably against Epic Games for using AI to replicate James Earl Jones' Darth Vader voice without union consultation, establishing important precedents for AI voice replication in entertainment. The intellectual property landscape was similarly transformed by Disney and Universal's landmark lawsuit against Midjourney for copyright infringement, alleging the AI platform enabled unauthorized reproduction of copyrighted characters. Internationally, regulatory momentum continued with Denmark proposing comprehensive legislation prohibiting unauthorized use of individuals' likenesses, while the platform-level response included Deezer's systematic flagging of AI-generated music content, with the streaming service detecting that 18% of daily uploads were completely AI-generated, prompting cuts to royalty payments for fraudulent AI content.



The 18% Problem

Streaming platform Deezer discovered that a staggering 18% of all songs uploaded daily were completely AI-generated.

Legal Landscape & Regulatory Response

The legal response to deepfake proliferation in Q2 2025 marked a watershed moment in regulatory action, with the signing of the Take It Down Act by President Trump representing the most significant federal legislation to date. This bipartisan bill mandates platform removal of non-consensual intimate imagery within 48 hours and establishes federal criminal penalties for distribution, addressing a critical gap that had allowed perpetrators to operate with impunity. At the state level, targeted legislation emerged from high-profile cases, with Florida's "Brooke's Law" requiring platform removal within 48 hours following a teenager's victimization, while Tennessee's "Preventing Deep Fake Images Act" criminalized malicious deepfake distribution as a felony response to meteorologist Bree Smith's harassment case.

Major Legislative Milestones - Q2 2025

May 19	**Take It Down Act (Federal)** President Trump signs bipartisan federal legislation criminalizing non-consensual deepfakes
April 15	**Florida's "Brooke's Law"** Mandates 48-hour platform removal following teenage victim case
March 26	**Tennessee Preventing Deep Fake Images Act** Makes malicious deepfake distribution a felony offense
February 12	**Louisiana Law 73.13** Criminalizes deepfakes depicting minors with 10-year minimum sentence

RESEMBLE AI

Major Deepfake Incidents Timeline

Apr 01

Non-consensual Explicit Content

Children Deepfakes

Deepfake technology was used to create non-consensual nude images of children in Malta and Spain. In Malta, at least four cases involved girls having their faces superimposed onto naked bodies or digitally undressed using AI-powered apps.

Apr 05

Political

Political Misinformation

China's state-run media used AI-generated videos featuring robots and consumers to criticize U.S. President Donald Trump and tariffs, highlighting the potential for high inflation and economic distress for Americans.

Apr 09

Political

Generative AI Videos

An AI-generated video mocking the idea of Americans working in factories, depicting them as depressed, obese, and dull-witted in sweatshop-like environments. The video was shared on TikTok and X, generating millions of views and sparking controversy.

Apr 14

Corporate Social Phishing

Zoom Social Phishing

A person posing to be the company's chief financial officer tricked the company's finance director into sending the funds after reaching out to them through the messaging app WhatsApp and setting up a Zoom call conference with the scammers.

RESEMBLE AI

Major Deepfake Incidents Timeline

Apr 21

Corporate Social Phishing

North Korean IT Workers

North Korean IT workers are using real-time deepfake technology to infiltrate organizations through remote work positions. They are using synthetic video feeds during job interviews to present different personas.

Apr 21

Political

Singapore Elections

A surge in AI-generated videos related to the Singapore GE2025 elections occurred after the polls were called. CNA detected 73 such videos posted between Apr 15 and Apr 19, with 11 containing digitally generated or manipulated visuals of prospective candidates.

Apr 22

Disinformation

Crosswalk Deepfakes

Crosswalks in Palo Alto, Menlo Park, Redwood City and Seattle were hacked to play satirical deepfakes of billionaire tech giants.

Apr 25

Disinformation

Legal Gen AI

Lawyers in the Coomer v. Lindell case submitted a brief containing nearly thirty defective citations, including fabricated cases, misquotes, and misattributed legal principles, generated by AI. The defense attorney admitted to using Gen AI

RESEMBLE AI

Major Deepfake Incidents Timeline

May 03

Misinformation

Donald Trump, the Pope

Donald Trump posted an AI-generated image of himself dressed as the Pope on social media during the Catholic mourning period for Pope Francis. The image was widely criticized as offensive and disrespectful by Catholics and others.

May 07

Corporate

Accenture CEO Spoofed

A deepfake video of Accenture CEO Julie Sweet was used in an attempt to defraud the company. The deepfake 'Julie' appeared on a video call with the company's finance head, instructing them to transfer funds.

May 15

Disinformation

South Asian Deepfake

Deepfake videos of PM Modi, Home Minister Amit Shah, and External Affairs Minister S Jaishankar were circulated with misleading claims by Pakistani users.

May 15

Social Engineering

Social Phishing

Hackers used AI-generated voice messages to impersonate senior US government officials in an attempt to break into their online accounts.

RESEMBLE AI

Major Deepfake Incidents Timeline

May 21

Legal

Gen AI in Court

A Toronto lawyer, Jisuh Lee, submitted a factum that included hallucinated cases generated by AI, misleading the court. She avoided contempt charges by taking responsibility and committing to remedial measures.

June 11

Legal

Midjourney Lawsuits

Disney and Universal are suing Midjourney for copyright infringement, claiming the AI image generator creates unauthorized copies of their copyrighted characters and trains its video service on copyrighted materials.

June 17

Celebrity

WWE Deepfake

A deepfake video surfaced online appearing to show Zoey Stark announcing her retirement from WWE, which caused concern among fans and her family. Stark later confirmed the video was AI-generated and that she is not retiring.

June 18

Corporate Social Phishing

Zoom call to install malware

North Korean BlueNoroff group deepfaked company executives in a Zoom call to trick an employee into installing macOS malware. The attackers contacted the target on Telegram, posed as professionals, and used a fake Zoom domain.

RESEMBLE AI

Major Deepfake Incidents Timeline

June 19

Political

Indian Plane Crash

AI-generated videos falsely depicting the aftermath of the Air India plane crash were circulated on Facebook. The videos were flagged as likely AI-generated by detection tools.

June 19

Financial Fraud

Investment Scheme

A deepfake video was created depicting President Anura Kumara Dissanayake endorsing a fraudulent investment scheme promising high returns.

June 20

Misinformation

Images of Military Jet

An AI-generated image falsely claimed to show a downed Israeli F-35 jet in Iran. The image, circulated online after Iranian state media claims, contained AI-generated distortions, mismatched symbols, and scale inconsistencies.

June 30

Social Media

Social Media Deepfakes

The South Dakota Department of Homeland Security (DHS), run by Kristi Noem, shared an AI-generated image on social media promoting a controversial immigration detention center called 'Alligator Alcatraz'.

Emerging Threats & Future Projections

The second quarter of 2025 revealed several alarming evolutionary patterns in deepfake technology that signal a fundamental shift toward more sophisticated and targeted attacks. Real-time deepfake capabilities emerged as a critical enterprise security threat, with North Korean IT workers successfully using live facial manipulation during remote job interviews to infiltrate major corporations and steal sensitive data. Simultaneously, the industrialization of AI-generated child sexual abuse material (CSAM) reached unprecedented scales, with law enforcement agencies across multiple countries reporting coordinated networks producing thousands of AI-generated images. The quarter also witnessed the emergence of "deepfake-as-a-service" operations, exemplified by Hong Kong-based syndicates that defrauded over HK$34 million using systematic deepfake romance scams across Southeast Asia, and the discovery that major streaming platforms were being flooded with AI-generated content—Deezer alone detecting that 18% of daily uploads were completely artificial.

Contact Us

detect@resemble.ai
Email

www.resemble.ai
Website

812 W Dana St, Mountain View, CA
Address

This report was compiled from publicly available information and represents the analysis of documented incidents. Due to the nature of deepfake activities, many incidents likely remain unreported, suggesting the actual scope may exceed what is documented here.