RESEMBLE AI

Q3 2025 DEEPFAKE REPORT

A Comprehensive Analysis of
Q3 2025 Deepfake Incidents

OCTOBER 2025

Executive Summary

The third quarter of 2025 witnessed an unprecedented surge in deepfake incidents, with 2,031 verified cases representing a significant evolution in both the scale and sophistication of synthetic media manipulation. This quarter's data reveals not just growth in volume, but a fundamental shift in the deepfake ecosystem toward industrialized, targeted attacks across multiple vectors.

2,031

Unique Incidents

The highest monthly total recorded to date.

48%

Celebrities Under Attack

targeted celebrities and public figures, with campaigns using their likenesses

385

Financial Devastation

Incidents reported direct financial losses, with sophisticated romance scams and investment frauds exploiting deepfake technology to build trust and credibility.

The Crisis Accelerates

The third quarter of 2025 represents a critical inflection point in the deepfake crisis. With 2,031 verified incidents, we're witnessing not just quantitative growth but qualitative evolution in attack sophistication.

2,031

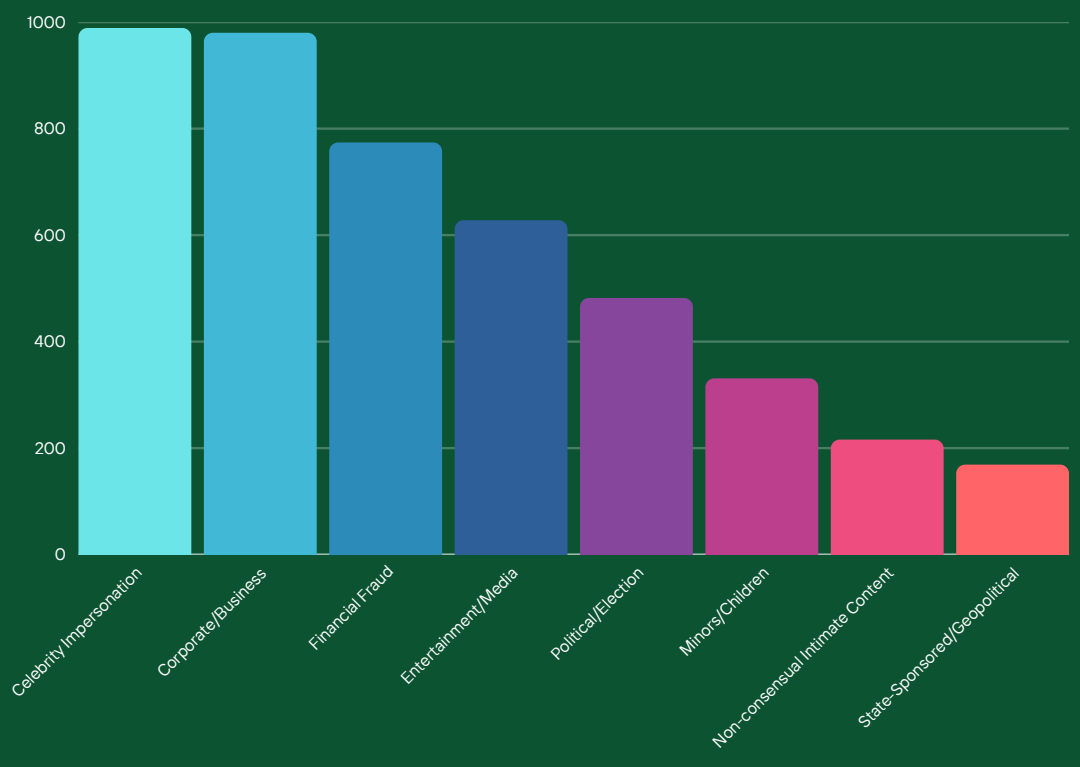
Total Unique Incidents in Q3 2025

The Multifaceted Attack Surface

RESEMBLE AI

Q3 2025 deepfake incidents span multiple attack vectors, with significant overlap between categories as criminals combine techniques for maximum impact. The data reveals sophisticated, multi-layered operations that blur traditional boundaries between fraud types.

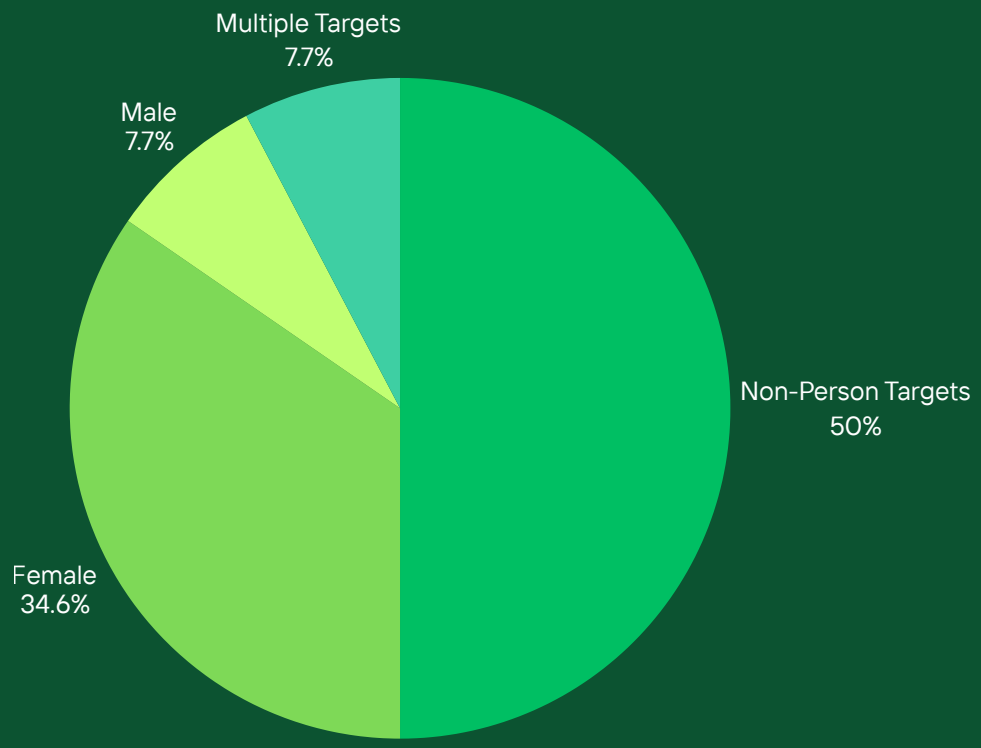
Number of Incidents by Category



Gender Dynamics and Digital Violence

Women are disproportionately targeted by deepfakes: Female targets outnumber male targets by 4.5:1 ratio (45 vs 10 incidents). This means that when deepfakes target specific individuals in Q3 2025, women are targeted 4.5 times more often than men.

Breakdown by Gender



Celebrity Industrial Complex Under Siege

Nearly half of all Q3 2025 deepfake incidents (48.7%) targeted celebrities and public figures. This represents not just individual attacks but a systematic exploitation of fame for criminal purposes. The celebrity deepfake economy has evolved into a sophisticated criminal enterprise, leveraging public trust in familiar faces to perpetrate fraud at scale.

Celebrity/Public Figure	Incidents
Will Smith	**24**
Barack Obama	**15**
Donald Trump	**13**
Marco Rubio	**10**
Alexandria Ocasio-Cortez	**7**
LeBron James	**6**
Taylor Swift	**5**

Corporate Infiltration and Business Warfare

The corporate sector faced an unprecedented assault in Q3 2025, with 48.3% of all incidents targeting businesses. The evolution from simple phishing to real-time deepfake impersonation during video conferences represents a quantum leap in social engineering sophistication.

980

Corporate/Business Incidents

Attack Vectors

- **Executive Impersonation:** Real-time deepfakes of CEOs and CFOs used during Zoom calls to authorize fraudulent wire transfers
- **Supply Chain Infiltration:** Deepfaked vendor representatives negotiating contract changes and payment redirections
- **Recruitment Fraud:** North Korean operatives using deepfakes to secure remote positions and plant malware
- **Investor Manipulation:** Fake executive announcements designed to manipulate stock prices
- **Brand Destruction:** Deepfaked product defects and false testimonials to damage competitor reputation

Singapore's September crisis demonstrates retail-level sophistication

RESEMBLE AI

13 Singaporeans lost more than SGD $360,000 to scammers impersonating telecommunications provider M1 Limited and the Monetary Authority of Singapore (MAS).

The attacks employed caller ID spoofing to display trusted brand names, voice deepfakes claiming to represent MAS financial authorities, and social engineering scripts alleging "suspicious activity" requiring immediate fund transfer to "safe accounts."

The urgency created panic conditions that bypassed victims' critical thinking, with fraudsters exploiting Singapore's high trust in governmental and telecommunications institutions.

The WhatsApp-Zoom Kill Chain

A particularly effective attack pattern emerged in Q3 2025

1

Initial contact via WhatsApp impersonating known executive

2

Urgent request for video conference citing 'confidential matter'

3

Real-time deepfake during Zoom call providing convincing visual confirmation

4

Authorization of fraudulent transaction or system access

5

Discovery only after significant damage inflicted

Financial Devastation

Financial fraud represented 38.1% of Q3 2025 incidents, with 385 cases reporting direct monetary losses. The sophistication of these attacks has evolved beyond simple impersonation to complex, multi-stage operations that build trust over weeks or months before striking.

774

Financial Fraud Incidents

The Romance Scam Evolution

Romance scams have transformed from text-based catfishing to sophisticated operations using real-time video deepfakes. Victims report months of video calls with their 'partners,' building deep emotional connections before the financial exploitation begins.

Cryptocurrency and DeFi Targeting

The intersection of deepfakes and cryptocurrency created particularly devastating losses. Deepfaked investment gurus, technical analysts, and even regulatory officials were used to promote fraudulent tokens, NFT projects, and DeFi protocols. The pseudonymous nature of crypto transactions made recovery virtually impossible.

Political Manipulation and Democratic Erosion

Q3 2025 witnessed 482 politically motivated deepfake incidents, representing 23.7% of all cases. These attacks targeted democratic processes across multiple nations, with sophisticated campaigns designed to manipulate public opinion, suppress voter turnout, and delegitimize election results.

482

Political/Election Related Incidents

Attack Patterns

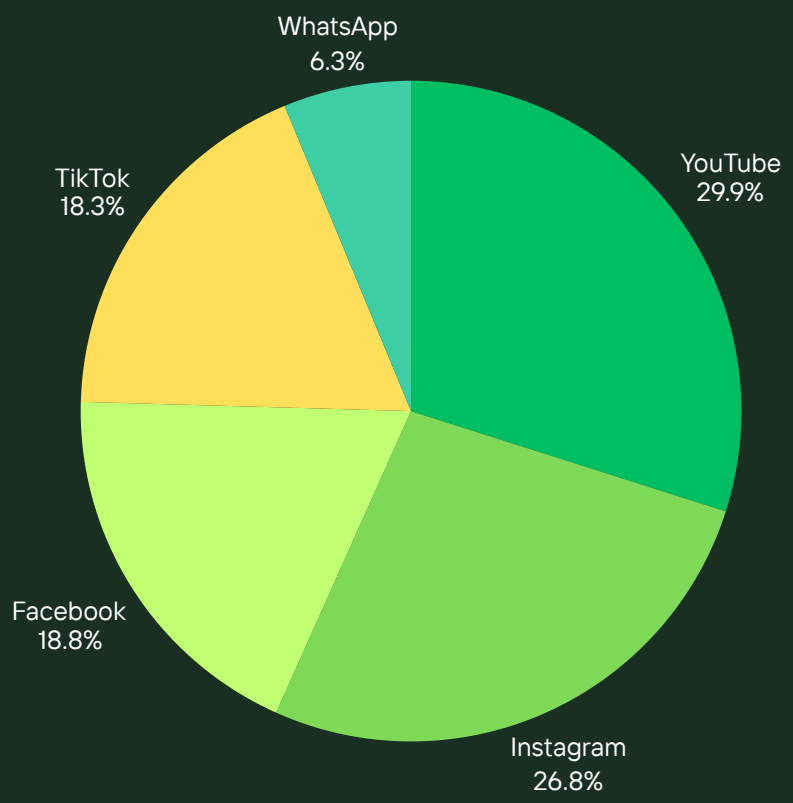
- **False Confessions:** Deepfaked videos of candidates admitting to crimes or corruption
- **Policy Reversals:** Fabricated statements showing candidates contradicting their platforms
- **Inflammatory Rhetoric:** Synthetic speeches containing racist or extremist language
- **Strategic Timing:** Release of deepfakes 24-48 hours before elections to maximize impact while minimizing fact-checking time

Platform Complicity and Technology Exploitation

RESEMBLE AI

Social media platforms emerged as both vectors and enablers of deepfake proliferation. Q3 2025 data reveals systematic failures in content moderation, with platforms profiting from advertisements for deepfake creation tools while simultaneously claiming to combat AI media abuse.

Platform Distribution



The Child Protection Crisis

The most alarming trend in Q3 2025 is the 331 incidents involving minors - representing 16.3% of all cases. This includes the creation of AI-generated child sexual abuse material (CSAM), cyberbullying through deepfaked content, and the targeting of educational institutions.

331

Unique Incidents

Incidents involving minors. AI-generated child sexual abuse material (CSAM)

Attack Patterns

- School-Based Attacks: Students creating deepfakes of classmates for bullying and harassment
- Predatory Behavior: Adults using deepfakes to groom and exploit minors
- Commercial Exploitation: Underground markets trading AI-generated CSAM at industrial scales
- Platform Failures: Social media platforms failing to detect and remove content involving minors

RESEMBLE AI

RESEMBLE AI

Major Incidents

July
01

Misinformation

AI Generated Flood

An AI-generated image was created using ChatGPT to depict the FOX43 studio flooded with 12 inches of water. The image was realistic, with reflections of lights in the water. The teleprompter and screen graphics in the image were jumbled and nonsensical, which is a telltale sign of an AI-generated image. This was used as a warning to highlight the potential for scams using AI-generated images of storm damage.

July
02

Law Enforcement

AI Altered Drug Evidence

Westbrook Police Department posted an AI-altered photo of drug evidence on social media. The photo was edited using ChatGPT to add the department's logo, which inadvertently changed other elements in the photo, such as removing stickers, altering lettering, and changing colors of items. Commenters noticed the discrepancies, leading to the department issuing an apology.

July
09

Political

Marco Rubio Deepfake

An attacker created a fake Signal account using the display name “Marco.Rubio@state.gov” and began contacting government officials with AI-generated voice messages that perfectly mimicked the Secretary of State’s voice and writing style. The imposter successfully reached at least five high-profile targets, including three foreign ministers, a U.S. governor, and a member of Congress, in an attempt to manipulate powerful government officials with the goal of gaining access to information or accounts.

RESEMBLE AI

RESEMBLE AI

Major Incidents

July 21

Misinformation

Fake Obama Arrest

Donald Trump shared a deepfake video on Truth Social depicting the arrest of Barack Obama by FBI agents in the Oval Office. The video shows Obama giving a speech, followed by his arrest while Trump watches. The video ends with Obama in a prison cell.

July 22

CSAM

Nudify App

Male students used "nudify" apps to create deepfake nudes of their female classmates. Incidents were reported in New Jersey, Texas, Washington, Florida, Pennsylvania, Southern California, and Iowa.

July 24

CSAM

Europol Child AI-Generated Images

A 29-year-old man in Denmark was found to have generated and sold thousands of AI-generated child sexual abuse images. He possessed over 36,000 such files and claimed to be among the top creators of realistic images of children. This led to a Europol operation involving 273 suspects and 25 arrests across 19 countries.

RESEMBLE AI

RESEMBLE AI

Major Incidents

August 17

Scam

AI Generated Books

AI-generated books imitating autobiographies of sports figures were created and sold on Amazon. The books were similar in style to the original autobiographies and were sold without clear warnings to consumers about their AI-generated nature. Steph Houghton, Neil Ruddock and Khalida Popal's autobiographies were imitated.

August 19

Political

Gavin Newsom & Kid Rock

Gavin Newsom posted an AI-generated image depicting Kid Rock, Tucker Carlson and Hulk Hogan supporting him, despite their actual political views. Kid Rock responded negatively.

August 20

Vishing

Qantas Vishing

AI voice cloning (vishing) used to impersonate a senior-level person to hack Qantas and to impersonate a company director to transfer $35 million.

August 21

Scam

Ohio Deepfake Video Scam

An Ohio man was scammed by a deepfake video impersonating Jelly Roll. The video claimed he'd won a $50,000 car, and he sent $70 in Apple gift cards for "shipping." The deepfake was detected because the driver's license in the video read "Jolly Roll" instead of "Jelly Roll."

RESEMBLE AI

RESEMBLE AI

Major Incidents

Sept
07

Scam

WhatsApp Deepfake

A 72-year-old homemaker in Hyderabad lost Rs 1.97 lakh to an AI voice cloning fraud. She received a WhatsApp message, seemingly from her sister-in-law, urgently requesting money. The voice on the other end sounded familiar, convincing her to transfer the money. Investigators believe fraudsters used AI to mimic her relative's voice.

Sept
24

Scam

Nirmala Sitharaman Financial Scam

Deepfake videos featuring Nirmala Sitharaman were circulated on social media, digitally manipulating her face and voice to make it appear as though she was personally endorsing a government-backed investment scheme promising unrealistic returns (deposit Rs 22,000 and earn Rs 50,000 every month).

Legal & Regulatory Landscape

Pending Federal Legislation

Several complementary bills remained under consideration in Q3 2025:

DEFIANCE Act: Re-introduced in May 2025 after the previous version passed the Senate in July 2024 but expired, this legislation would provide victims of non-consensual sexual deepfakes with a federal civil cause of action and statutory damages up to $250,000.

NO FAKES Act: Re-introduced in April 2025, this bipartisan legislation would create a federal right of publicity specifically for digital replicas, protecting individuals from unauthorized use of their likeness or voice. The Act would not preempt state laws existing as of January 2, 2025 or state statutes specifically addressing sexual or election-related deepfakes.

State-Level Developments

Pennsylvania Act 35
On July 7, 2025, Pennsylvania Governor Josh Shapiro signed Act 35 (formerly Senate Bill 649) into law, establishing criminal penalties for creating or disseminating deepfakes with fraudulent or injurious intent. The legislation includes:

Covered Conduct: Criminalization of creating, disseminating, or facilitating the creation of forged digital likenesses when a party reasonably should have known the material was fake.

Protected Expression Carve-Outs: Exemptions for satire, content in the public interest, and technology companies providing the means to create deepfakes without intentional facilitation of harmful use.

Disclaimer Defense: Placing a clear disclaimer on digital content identifying it as fake constitutes a legal defense.

Context: The law builds upon earlier Pennsylvania statutes criminalizing non-consensual distribution of AI-generated sexual imagery and follows enforcement activity involving deepfake content depicting minors.

Washington State House Bill 1205
Effective July 27, 2025, Washington State enacted comprehensive deepfake legislation criminalizing the intentional use of "forged digital likeness" with intent to defraud, harass, or injure. The statute applies to synthetic audio, video, and images created through AI manipulation.

The state-level momentum reflects what legal scholars characterize as "hundreds of active proposals" with approximately 25 deepfake bills introduced weekly across U.S. state legislatures during 2025.

Broader State Trends

By Q3 2025, the regulatory landscape included:

- 39 states with laws regulating intimate or erotic deepfakes, including CSAM and revenge porn
- 25 states with election-related deepfake regulations requiring disclaimers on AI-altered political content
- Over 30 states with proposed or enacted political deepfake laws mandating watermarks or disclosures

Notable Q3 2025 Legal Cases and Incidents

Voice Cloning Litigation
Aero Cosmetics v. Schmitman (September 2025): Vehicle wax company filed preemptive lawsuit after rogue employee used AI to clone voice actor Craig Schmitman's voice without authorization. The company withdrew all AI-generated content and issued multiple apologies, highlighting corporate liability risks for unauthorized AI deployments by employees.

Asha Bhosle Personality Rights (October 2025): Though technically early Q4, the Bombay High Court granted legendary singer Asha Bhosle ad-interim protection against AI voice cloning and image misuse. Justice Arif Doctor ruled that making AI tools available to convert voices into celebrity likenesses without permission violates personality rights. Amazon, Flipkart, and Google were directed to remove infringing content.

Attorney Sanctions
Arizona Federal Court (August 14, 2025): Judge revoked attorney Maren Bam's pro hac vice status and imposed multiple sanctions after discovering the majority of legal citations were AI-generated fabrications. The case underscores judicial intolerance for AI-generated false legal authority.

Corporate Actions
Vogue/Guess AI Model Controversy (August 2025): American Vogue's August 2025 issue featured Guess advertisements with AI-generated models, sparking widespread reader backlash, subscription cancellations, and calls for boycotts. The incident highlighted ethical concerns beyond legal compliance, particularly regarding representation, diversity, and job displacement for human models.

International Regulatory Developments

China: Comprehensive AI Content Labeling Framework

September 1, 2025, marked the implementation of China's groundbreaking "Measures for Labeling Artificial Intelligence-Generated Content" alongside mandatory national standard GB 45438-2025. Released by the Cyberspace Administration of China (CAC) on March 14, 2025, these regulations represent the world's most comprehensive AI content labeling framework.

Dual Labeling System:
Explicit Labels: Visible indicators (text, audio, or graphics) clearly informing users when content is AI-generated. Required on all AI-generated text, images, audio, video, and virtual scenes.
Implicit Labels: Embedded metadata including "AIGC" markers, content producer identity (unified social credit code or citizen ID), and unique content identifiers from AI service providers.

Platform Obligations:
Content distribution platforms must implement detection mechanisms categorizing AI-generated content into three tiers:

- Confirmed AI-Generated: Detected implicit labels trigger mandatory explicit labeling
- Possible AI-Generated: User reports without implicit labels require "possibly AI-generated" warnings
- Suspected AI-Generated: Evidence of AI generation without labels triggers "suspected AI-generated" notices

Traceability Requirements:
Providers must retain generation logs and ensure full traceability through metadata and content IDs, enabling regulators to trace content to its origin.

Enforcement Mechanisms:
The CAC's 2025 "Qinglang" enforcement campaigns specifically target AI-generated content violations. In Q2 2025 alone, over 3,500 problematic AI products were handled, more than 960,000 pieces of illegal content removed, and approximately 3,700 accounts addressed.

China's framework is regarded as setting international precedent. As noted by legal analysts, "China's law introduces the most comprehensive requirements to date for the labeling, traceability, and accountability of AI-generated content anywhere in the world."

Conclusion: systemic vulnerabilities demand architectural transformation

Resemble AI's Multimodal Deepfake Detection model, built on multimodal analysis and continuous retraining architectures, addresses precisely what the threat landscape now demands – real-time detection capabilities that achieve 94-96% accuracy through ensemble methods, behavioral biometrics, and daily model updates that adapt to evolving attack vectors.

The quarter's developments validate Resemble AI's strategic pivot toward enterprise deepfake detection. State-sponsored infiltration programs, real-time interactive impersonation during live video calls, and coordinated multi-channel attacks demonstrate that deepfake defense is no longer optional security enhancement but critical infrastructure for digital trust. As regulatory frameworks like the EU's mandatory labeling requirements create compliance obligations, and as Gartner projects that 30% of enterprises will find standalone identity verification unreliable by 2026, Resemble AI's detection-first approach positions the company as essential partner for financial services, government agencies, and enterprises facing existential authentication challenges. The crisis has arrived, and with it, unprecedented demand for the trust layer that Resemble AI provides.

RESEMBLE AI

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.