WHITE PAPER

Synthetic AI

Deepfakes, Fake Proof, and the Brand Defense Defense Playbook

A 5-part article series distilled into an executive brand defense framework

---

Logan Sivanasen | March 2026

logansivanasen.com

NAVIGATION

Table of Contents

This white paper distills five original LinkedIn articles into a single executive operating framework for brand defense in the Synthetic AI era. It covers the core risk model, five-chapter playbooks, three real-world case studies, and a practical summary with resources.

Five-Chapter Playbook

Three Real-World Case Studies

Also Inside

CORE FRAMEWORK

Synthetic AI Turned Trust Into an Operating Operating System Problem

The old model assumed proof was stable and impersonation was rare. That model is obsolete. The conditions that made it work slow-moving synthetic tools, high production costs, and audiences with time to verify, no longer exist. Today, fake proof is scalable, cloned identities are convincing at scroll speed, and synthetic incidents can reach millions of people before a single internal alert fires.

Most teams are still treating this as a PR problem, a legal problem, or a cyber problem. It is all three, running simultaneously, compounding each other, and moving faster than any siloed team can respond. A PR team without legal authority cannot contain it. A legal team without brand context cannot message it. A cyber team without comms support cannot explain it. The silos are the vulnerability.

The fix is not a new policy document or a one-off training session. It is a shared operating model; one that connects marketing, comms, legal, security, support, and paid media into a single response architecture. That architecture must be designed, tested, and owned before the incident arrives. Because when it does, the window to build it has already closed.

The Shift

Generated content scales faster than verification. The asymmetry is structural, not temporary.

The Risk

Trust erosion, demand theft, fake proof, fraud, and public confusion: these often run simultaneously.

The Failure Mode

Slow detection, scattered proof, weak escalation, and no single source of truth.

The Fix

Clear ownership, proof controls, hijack monitoring, risk routing, and response choreography.

This framework helps leaders move from reactive panic to operational control. Synthetic AI is no longer a side issue. It is now a brand, trust, proof, demand, and response problem.

CHAPTER 1

Your Brand Will Get Impersonated. What To Do First.

The first loss in a synthetic incident is not always money. It is control. Impersonation now looks real enough to pass scroll speed. Marketing owns more attack surface than most leaders admit: every public-facing executive, every branded channel, every verified handle is a potential target. The first 30 minutes matter more than any long policy document. Teams need one incident owner, one source-of-truth link, and one repeatable loop.

Impersonation Now Looks Good Enough

Synthetic quality has crossed the threshold of casual scrutiny. Audiences cannot reliably self-detect fakes at scroll speed.

Detection Time Is a Trust Metric

Every hour an impersonation runs unchallenged is an hour your audience is being misled under your brand.

Triage Like a CISO, Speak Speak Like a CMO

Incident response needs security-grade speed and brand-grade clarity. Both matter simultaneously.

The Brand Defense Loop

Map

Monitor

Move

Message

First 30 Minutes Checklist

01

Freeze outgoing changes

02

Open incident channel

03

Assign incident lead

04

Capture evidence

05

Publish one source of truth

Read the original article

CHAPTER 2

Fake Proof Will Hit Your Campaigns. Build The The Control Plan.

Synthetic AI does not only impersonate brands; it fabricates evidence. Fake screenshots, fake dashboards, fake testimonials, and fake case signals are a direct GTM risk. Real proof is expensive and slow. Fake proof is cheap and fast. That asymmetry is the threat. Proof must be treated like a regulated asset inside go-to-market. Tooling helps, but discipline wins.

Proof Is Now a Governed Governed Asset

If it is market-facing, it needs controls. Proof without provenance is a liability waiting to surface.

Attackers Only Need to to Hack Your Narrative Narrative

They do not need to breach your systems. A convincing fake claim can do more damage than a data breach.

Detection Without Ownership Is Only a Smoke Alarm

Knowing a fake exists means nothing without a named owner who can act on it immediately.

The Proof Control Plan

Ownership

Named Proof Steward across Marketing, RevOps, Risk, Legal, Finance, and Security

Standards

Defined formats, source requirements, and consent documentation

Pre-Flight Checks

Verification gate before any proof asset goes market-facing

Response Playbook Playbook

Named escalation path for disputed or fabricated proof claims

Proof KPIs to Track

- **Approved proof usage rate**
- **Traceable consent rate**
- **Post-launch correction rate rate**
- **Time to proof approval**
- **Time to public response on on disputed claims**

Read the original article · **Reference:** FTC crackdown on deceptive AI claims

CHAPTER 3

Lookalike Ads and Brand Hijacks. Stop Demand Theft.

Brand hijacks are not a vague brand safety issue: they are measurable revenue leaks. Lookalike ads, rogue affiliates, grey-market resellers, and deceptive brand bidding sit on your demand and redirect it. Platforms enforce policy, not your revenue threshold. Hijacks hide inside normal performance noise: CPC, CPA, refunds, and NPS signals. This is a GTM control issue, not only a paid media issue.

Brand Terms Are Now an Attack Surface

Every branded search term is a bidding opportunity for bad actors. Your brand equity funds their acquisition.

Hijacks Hide Inside Performance Noise

CPC spikes, CPA drift, refund trends, and NPS dips are often the first signal, not a platform alert.

Demand Theft Damages Damages Trust and Data Data Simultaneously

Hijacked journeys corrupt attribution, inflate acquisition costs, and erode audience trust in parallel.

The Brand Hijack Control Loop

1. **Map hijackable journeys**
2. **Monitor like revenue**
3. **Classify hijacks**
4. **Pull the right levers**
5. **Feed learning back into strategy**

Read the original article · **Reference:** AI Time Journal: The Silent Killer: Competitors Stealing Your Brand in Google Ads

Competitors and bad actors use geo-targeting, cloaking, multi-step redirects, and AI bots to evade detection inside affiliate and paid search channels. Traditional manual monitoring cannot keep pace. Automated tools such as Adthena and BrandVerity provide continuous global coverage.

CHAPTER 4

Ship Faster Without a PR Crisis. Two Lanes, One Log.

Synthetic AI changed the burden of proof. Speed without traceability becomes reputational debt. Risk is not evenly distributed across content: some assets carry far more exposure than others. Teams fail when everything sits in one approval queue. Governance needs visibility and traceability, not theater. The answer is not slowing down; it is routing correctly.

Risk Is Uneven Across Content

Not every asset needs the same gate. Treating all content equally creates bottlenecks without reducing real risk.

Synthetic AI Raises the Proof Burden

Audiences and regulators now expect traceability. Claims without evidence trails are increasingly indefensible.

You Cannot Govern What What You Cannot See

A shared log is not bureaucracy. It is the only way to know what shipped, who approved it, and what evidence backs it.

Two Lanes, One Log

Lane A: Fast Lane

Low risk; High throughput; Streamlined approval

Lane B: Controlled Lane Lane

High risk; High defensibility; Full evidence trail

One shared log for traceability, ownership, and evidence across both lanes.

The Evidence Ladder

Level 1

Source link

Level 2

Source + excerpt + date

Level 3

Source + calculation trail

Level 4

Source + methodology + owner sign-off

Read the original article · **Reference:** McKinsey: Speed and Safety in GenAI

CHAPTER 5

The 24-Hour Response System

Most teams do not fail on intelligence: they fail on choreography. The first 24 hours should run on three parallel tracks, not one sequential chain. Severity routing protects both speed and judgment. Cadence beats silence and over-explanation. You do not win the first day by proving everything. You win it by coordinating everything.

Synthetic Incidents Spread Faster Than Verification Cycles

Waiting for full proof before communicating cedes the narrative to the incident itself.

Run Takedowns, Comms, and Recovery in Parallel

Sequential response is too slow. Three simultaneous tracks are the minimum viable structure.

Speed Is Not the Danger. Uncoordinated Speed Is.

Fast and fragmented is worse than slow and coordinated. The log and the owner make speed safe.

The 24-Hour Response Card

0–2 hrs

Contain: Assign owner; open incident channel; capture evidence; freeze outgoing.

2–8 hrs

Clarify: First holding statement; takedown requests; stakeholder brief; source-of-truth page live.

8–24 hrs

Stabilize: Monitor mirrors; support ticket review; decision log updated; next update scheduled.

Response KPIs

- **Time to incident owner**
- **Time to first takedown**
- **Time to first holding statement**
- **Mirror rate**
- **30-day control closure rate**

Three Parallel Tracks

Takedowns

Platform reports, legal notices, affiliate flags

Comms

Holding statement, internal brief, media monitoring

Recovery

Source-of-truth page, support triage, trust signals

Read the original article · **Reference:** NIST AI Risk Management Framework

CASE STUDY 1 | PART 1 OF 2

Arup and the US$25M Deepfake Fraud

Synthetic AI does not need to breach your stack. It only needs to look credible.

What Happened

A finance employee at Arup's Hong Kong office joined what appeared to be a legitimate executive video call. Every person on screen was a deepfake. HK$200 million was transferred across 15 transactions to 5 bank accounts before the fraud was discovered. Arup confirmed no internal systems were compromised.

How the Attack Unfolded

01

Identities Cloned

Public executive video and audio used to build synthetic personas.

02

Video Call Staged

Teams-style call with multiple synthetic participants on screen.

03

Instruction Issued

Employee acted on a direct financial transfer instruction.

04

HK$200M Transferred

15 transfers to 5 accounts before any check triggered.

HK$200M	US$25M	15	5
Total Lost	Equivalent	Transfers	Bank Accounts

> "The employee followed normal process. The process had been hijacked at the human layer, not the technical layer."

CASE STUDY 1 | PART 2 OF 2

Arup and the US$25M Deepfake Fraud

Readiness matters before the incident, not during it.

Why It Matters

The Arup fraud proves synthetic AI has moved from reputational threat to operational risk. No technical breach needed. Just a convincing performance and a workflow with no out-of-band verification. Every organisation using video calls for financial authorisation has the same exposure.

The Verification Gap

When the call looks real, the voice sounds real, and the faces match, the only defense is a protocol outside the compromised channel. That protocol must exist before the incident. The window between instruction and transfer is too short to build it during.

What to Build Now

Out-of-Band Verification

Second confirmation channel for any financial instruction above a defined threshold.

Video Call Identity Policy

Video presence is insufficient proof of identity. Policy, not guidance.

Deepfake Scenario Exercises

Tabletop exercises with finance and operations teams before an incident.

Named Incident Owner

Ownership assigned before the call comes in, not after.

The Broader Implication

Video Workflows Are Exposed Exposed

Any org authorising actions over video calls needs its verification layer redesigned now.

Executive Profiles Are Attack Material

Every public video and interview is training data for synthetic impersonation.

Cyber Cannot Own This Alone Alone

Finance, operations, and leadership must share ownership.

Lesson: Synthetic AI does not need to breach your stack first. It only needs to look credible enough to break your workflow.

Source: BBC News: Arup deepfake fraud · https://www.bbc.com/news/business-68614677

CASE STUDY 2 PART 1 OF 2

Celebrity Deepfake Investment Scams on Meta in Australia

This is what synthetic brand hijack looks like at platform scale.

What Happened

Between January and August 2024, Australians reported A$43.4 million in losses from social media scams, with ~A$30 million tied to fake investment scams. Meta's FIRE program with seven Australian banks removed 9,000+ scam pages and 8,000 AI-generated celebrity investment ads. One victim lost A$80,000 after engaging with a deepfake Elon Musk investment video.

The Scam Pipeline

01

Likeness Cloned

Public celebrity content used to generate synthetic endorsement ads.

02

Ads Placed on Meta

Distributed as paid advertising on Facebook and Instagram.

03

Victims Funnelled

Audiences directed to fake investment platforms.

04

Funds Lost

Financial damage done before takedowns were confirmed.

A$43.4M
Reported Losses

~A$30M
Investment Scams

9,000+
Pages Removed

8,000
AI Celebrity Ads

7
Banks Involved

A$80K
Victim Loss

> "At scroll speed, audiences had no reliable signal to distinguish a synthetic celebrity endorsement from an authentic one."

CASE STUDY 2 | PART 2 OF 2

Celebrity Deepfake Investment Scams on Meta in Australia

Platform enforcement operates on policy timelines, not your trust thresholds.

Why It Matters

This case is not about a rogue post. It is a commercialised pipeline at platform scale: synthetic face, paid placement, and measurable financial damage. The scam ads used the same visual format as legitimate sponsored content. By the time reporting begins, the audience has already been reached.

The FIRE Program Response

17,000+ pieces removed

Meta and seven Australian banks removed 9,000+ scam pages and 8,000 AI celebrity investment ads.

Reactive, not preventive

Cross-sector coordination was significant. But it operated after the damage was done.

ACCC flagged systemic risk

Highlighted as platform-scale fraud risk, not isolated misuse.

What to Build Now

Likeness Monitoring

Continuous monitoring for unauthorised use across Meta, TikTok, and YouTube.

Takedown Templates Templates

Pre-drafted requests ready before an incident. Speed over perfection.

Legal Escalation Paths Paths

Independent of platform response timelines.

Audience Alert Protocols

Define the threshold to alert your audience directly.

Cross-Functional Ownership

Brand, legal, and paid media must share accountability.

The Broader Signal

Australia Is Not Isolated

The same playbook runs in Brazil, the UK, and Southeast Asia. The tools and format are identical.

Paid Distribution Amplifies Damage Damage

Paid ads reach targeted audiences at scale within hours. Damage compounds faster than reactive workflows can match.

Public Profiles Are Permanent Risk

If your brand or executives have public-facing profiles, the infrastructure to exploit them already exists.

Lesson: Fake proof now operates as an acquisition engine for fraud. If your audience recognises the face, tone, and format, the scam already has a head start.

Source: ACCC Scamwatch and Meta FIRE Program · https://www.accc.gov.au/media-release/australians-lose-43-million-to-social-media-scams

CASE STUDY 3 | PART 1 OF 3

Gisele Bündchen Deepfakes in Instagram Scam Ads

This is not fringe misuse. It is commercialised face theft tied to paid distribution and measurable fraud.

What Happened

Brazilian investigators uncovered a scam ring using Instagram ads featuring deepfakes of Gisele Bündchen to push fraudulent offers. Authorities identified 20M+ reais (~US$3.9M) in suspicious funds, arrested 4 suspects, and froze assets across 5 states via paid Instagram ad distribution.

How the Scam Operated

01

Likeness Cloned

Celebrity imagery generated synthetic endorsement content.

02

Ads Placed on Instagram

Distributed as legitimate sponsored posts.

03

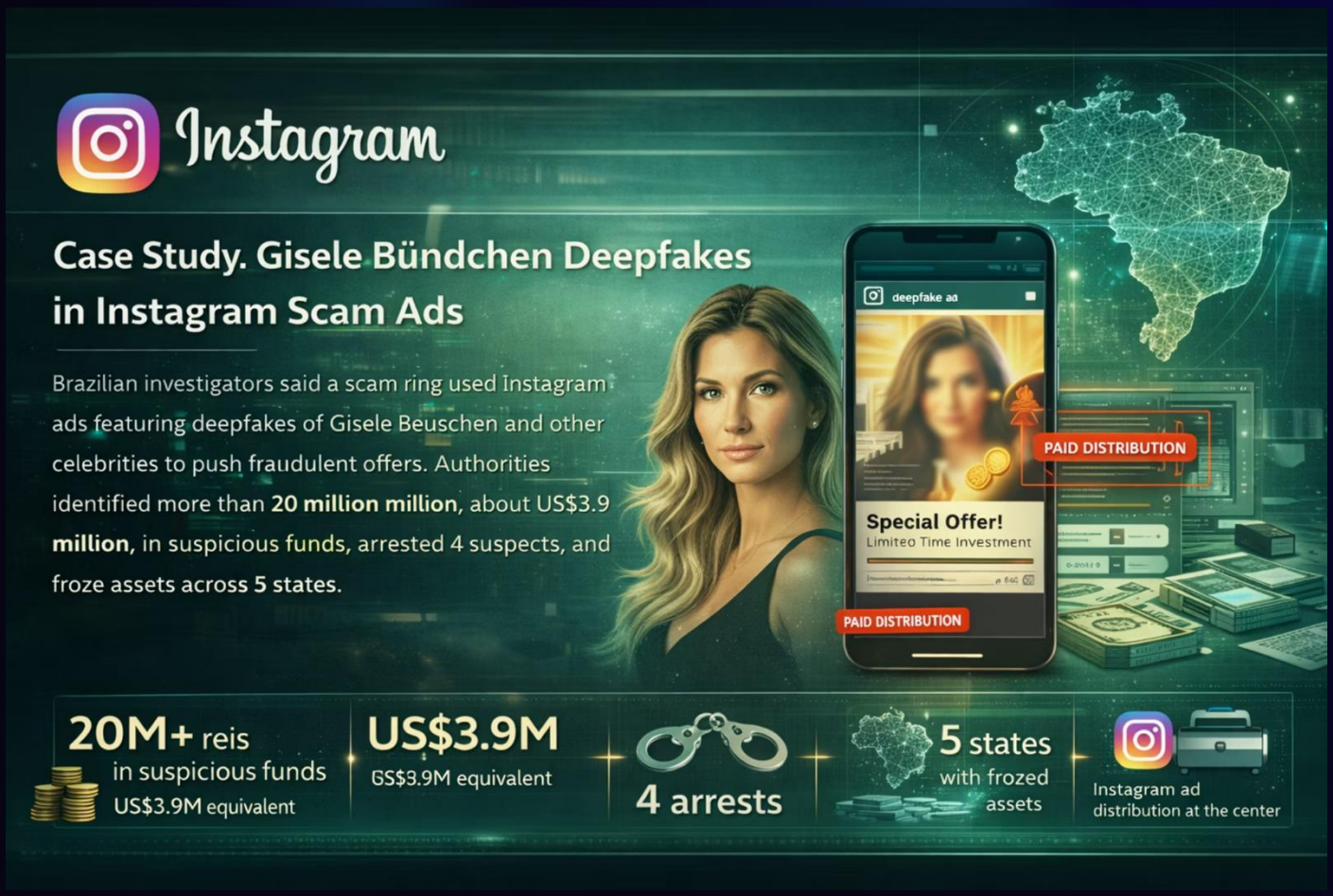
Victims Funnelled

Audiences directed to fake investment platforms.

04

Investigators Intervened

Police traced funds, arrested 4, froze assets.



20M+

Reais Suspicious Funds

US$3.9M

Equivalent Value

4

Arrests Made

5

States, Assets Frozen

Meta/IG

Distribution Channel

> "The scam ads used the visual grammar of legitimate advertising. By the time reporting begins, the audience has already been reached."

CASE STUDY 3 | PART 2 OF 3

Gisele Bündchen Deepfakes in Instagram Scam Ads

If your likeness, format, and trust cues are easy to mimic, scam ads get a head start before reporting even begins.

Why It Matters

This is a commercialised pipeline: synthetic face, paid placement, and measurable fraud at scale. The scam ads used the visual grammar of legitimate advertising. The platform's distribution engine was the mechanism. By the time reporting begins, the audience has already been reached.

The Broader Signal

Same playbook runs globally

Documented in Australia, UK, US, and Southeast Asia. Identical tools, format, and distribution.

Law enforcement is not a fast channel

The Brazil investigation took months. Criminal prosecution cannot be the primary defense.

Instagram was the crime scene, not the cause

Ad tools were exploited, not compromised. Brands must assess their own exposure accordingly.

What Brand and Legal Teams Should Build Now

Continuous Likeness Monitoring

Paid social monitoring for unauthorised use across Meta, TikTok, and YouTube.

Pre-Drafted Takedown Takedown Templates Templates

Ready before an incident. Speed over perfection in the first hour.

Legal Escalation Paths

Independent of platform response timelines.

Audience Alert Protocols

Define the threshold to alert your audience directly.

Cross-Functional Ownership

Brand, legal, and paid media must share accountability.

The Broader Implication

Public Figures Are Permanent Attack Attack Surfaces

Every public appearance and media asset is source material for synthetic impersonation.

Paid Distribution Amplifies Damage

Ads reach targeted audiences at scale within hours. Damage compounds faster than reactive workflows.

Law Enforcement Is Not a Fast Channel

Thorough but slow. Brands cannot rely on criminal prosecution as primary defense.

Lesson: Fake proof now operates as an acquisition engine for fraud. If your likeness, format, and trust cues are easy to mimic, scam ads get a head start before reporting even begins.

Source: Brazilian Federal Police and Reuters: Gisele Bündchen deepfake Instagram scam

SERIES SUMMARY

Brand Defense in the Synthetic AI Era

This series shows that Synthetic AI changed the nature of brand defense. What used to sit across PR, paid media, content review, legal, and cyber now needs one shared operating model. The leaders who win will not be those with the loudest AI strategy. They will be the ones with the clearest proof, fastest routing, strongest source of truth, and calmest response under pressure.

01

Impersonation is now operational risk, not edge-case noise

02

Proof needs governance before before campaigns go live

03

Brand hijacks are measurable demand theft

04

Speed needs risk routing and traceability

05

The first 24 hours require choreography, not panic

Five Non-Negotiables

One Owner

Every incident, every proof asset, every hijack needs a named human accountable for it.

One Proof Standard

Consistent evidence requirements across all market-facing claims, pre-launch.

One Hijack Control Loop

Map, monitor, classify, act, and feed learning back: continuously.

One Risk-Routing Log

Shared visibility across marketing, legal, security, and comms.

One Source-of-Truth Page

Ready to publish within the first two hours of any synthetic incident.

Links to All 5 Original Articles

Chapter 1: Your Brand Brand Will Get Impersonated. What To What To Do First.

Chapter 2: Fake Proof Will Hit Your Campaigns. Build The Control Plan.

Chapter 3: Lookalike Ads and Brand Hijacks. Hijacks. Stop Demand Demand Theft.

Chapter 4: Ship Faster Without a PR Crisis. Two Lanes, One Log.

Chapter 5: The 24-Hour Hour Response System. System.

REFERENCES AND RESOURCES

Research, References and Starter Kits

Research and Reference Links

- NIST: Synthetic Content Overview
- NIST AI Risk Management Framework
- FTC: Crackdown on Deceptive AI Claims
- McKinsey: State of AI 2024
- McKinsey: Why Digital Trust Truly Matters
- McKinsey: GenAI Speed and Safety
- Reuters: Stronger Deepfake Detection Standards
- Bird & Bird: EU AI Act Transparency Code
- LinkedIn: Fake Profile Reporting
- YouTube: Synthetic Content Disclosure

Download the Synthetic AI Brand Defense Starter Kit

Everything you need to operationalize your brand defense framework immediately:

1. First 30 Minutes Checklist
2. Proof Control Plan Canvas
3. Brand Hijack Control Loop
4. Two Lanes, One Log Template
5. 24-Hour Response Card