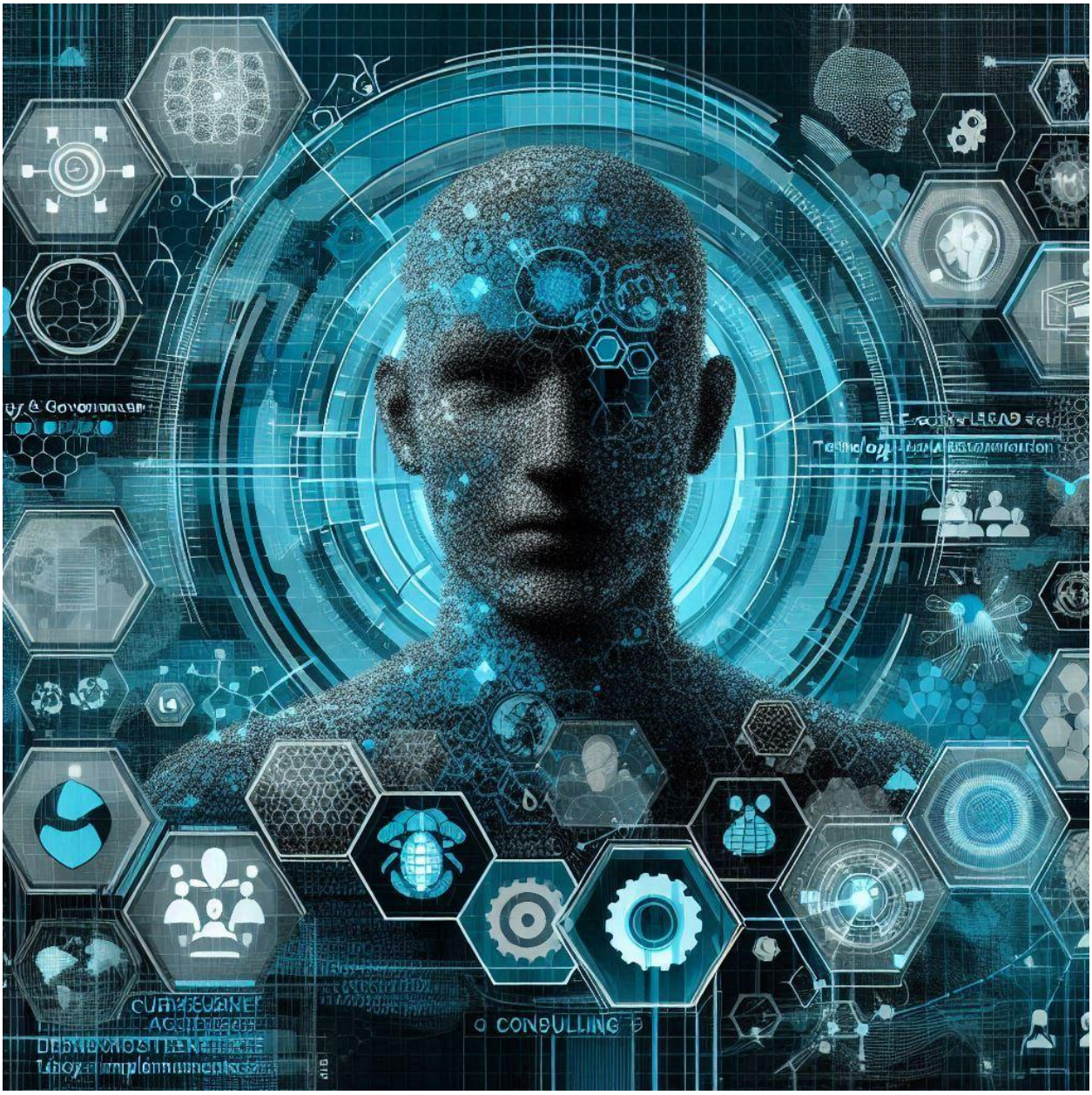

Trends in Effectiveness of AI-Enabled Social Engineering

Executive Summary

Over the past two decades (2000-2025), the effectiveness of social engineering attacks (phishing, impersonation, deepfakes, other manipulation) has substantially increased. This growth is particularly marked since roughly 2018–2020, coinciding with the rise of accessible generative AI, voice-/video-cloning tools, and automation of attacker workflows.

Key drivers of this trend include:

- The increased availability of personal data (social media, breaches) enabling better targeting.
- Automation and scaling via AI (bots, synthetic media) reducing attacker cost and friction in operations.
- Enhanced realism in impersonation (voice, video deepfakes) that undermines traditional human cognitive defenses.
- The "economics of fraud" shifting: lower cost per attempt, higher volume, and more high-value successes (e.g., corporate wire fraud).
- A lag in corresponding defense improvements (user training, detection technologies, process redesign) leaving soft targets.

This report charts the development in phases, quantifies indications of rising effectiveness, describes enabling technologies, highlights case-studies, and finishes with implications for defenders and predictions going forward.

1. Historical Phases in Social Engineering (2000–2025)

Phase I: 2000–2010... Classic Phishing & Opportunistic Attacks

- The early 2000s saw widespread "spray-and-pray" phishing: generic emails, spoofed domains, credential theft.
- Social engineering mostly human-driven, with manual email creation, simple impersonation, limited personalization.
- Data on success rates is sparse, but indicators show increasing volume of attacks (for example, phishing reports from APWG). Wikipedia
- Defense was primarily signature-based filters, user awareness training, and basic domain controls.

Phase II: 2010–2018... Targeted & Spear-Phishing Emergence

- Attackers began using more targeted approaches: spear-phishing emails referencing specific persons/projects, more social-engineering content.
- The growth of social media, breach-data exposure, and profiling enabled more effective targeting.
- While not yet fully driven by generative-AI, the tools improved attacker toolkit (automated email senders, phishing kits).
- Attackers began targeting C-suite executives, wires, business-email-compromise (BEC).
- Defenders improved controls (MFA, endpoint detection), but attacker sophistication also rose.

Phase III: 2018–Present... AI-Enabled Social Engineering & Deepfakes

- The advent and maturation of generative AI (text, voice, video) drastically changed the attacker economics.
- Examples: AI bots generating very credible phishing emails, voice-cloning for vishing, video deepfakes for impersonation of executives. SpringerLink
- Automation enables thousands of personalized outreach attempts, voice or video impersonation, richer social-engineering narratives. CrowdStrike

- For instance, one report noted that deepfake-enabled vishing surged by 1,600 % in Q1 2025 relative to end-2024. Right-Hand Cybersecurity
- The Covid-19 pandemic accelerated remote working, video-calls, digital communications... increasing the attack surface and opportunities for impersonation. arXiv

2. Indicators of Increased Effectiveness

Volume & Targeting

- According to a 2024 report by Allot Communications: a 54 % increase in phishing attacks in 2024 compared with prior periods (based on network-protection activity). Allot
- Deepfake/biometric fraud: one blog reports "deepfakes already account for 40 % of all biometric fraud" and that in Australia 20 % of companies reported deepfake scams in the past year. ZeroThreat
- 61 % of organizations saw an increase in deepfake attacks in the past year. Cobalt

Success / Impact

- Large-scale, high-value incidents: e.g., a finance worker at a multinational firm was tricked into paying $25 million after deepfake impersonation of a CFO. SpringerLink
- Forecasts: **Deloitte predicts** U.S. losses from generative-AI fraud will reach **US$40 billion by 2027**. appgate.com

Attack Quality & Detection Difficulty

- Traditional cues (grammatical errors, generic greeting) are disappearing: "AI-generated communications nearly indistinguishable from legitimate messages." Paubox
- Attacks are longer, **more context-rich**: for example, phishing emails are three times longer in 2023 than in 2021. Help Net Security
- Voice and video impersonation enable bypass of biometric or interpersonal verification: e.g., regulatory guidance notes that AI-driven social engineering attacks have led to unauthorized actions by mimicking appearance/voice. Department of Financial Services
- Automation and scale: "AI tools can now conduct thousands of phone calls simultaneously, each highly personalized." CrowdStrike

Shift in Methods

- Rise of phishing-as-a-service (PhaaS) and polymorphic phishing campaigns: Allot report mentions "AI-powered polymorphic phishing campaigns" gaining traction. KnowBe4
- Deepfake-driven social engineering moving into voice and video domains, not just email. MDPI

3. Enabling Technologies & Mechanisms

Generative Text & Emails

- Large Language Models (LLMs) enable crafting of highly context-aware, personalized phishing content referencing specific persons, projects, using correct tone, grammar and style. arXiv
- Bots can mine social media data, publicly-exposed breached records, communication style signals to create more believable lures. Yale Cybersecurity

Voice & Video Cloning (Deepfakes)

- AI-driven voice synthesis and deepfake video enable impersonation of real individuals in voice calls or video conferences: "voice phishing (vishing) and deepfake attacks are rising as attackers use generative AI." SpringerLink
- Regulatory notes highlight that threat actors have used deepfakes to defeat biometric verification and call centre checks. Department of Financial Services

Automation / Scale / Service Platforms

- Phishing-as-a-service (PhaaS) platforms combine templates, automation, and infrastructure to scale attacks. Allot
- Attackers increasingly treat social engineering like a production process: mass-data collection → segmentation → dynamic content generation → multichannel delivery (email, voice, deepfake video). World Economic Forum

Social Media & Remote Work Amplification

- The shift to remote work, video conferencing, chat platforms has enlarged the "human interface" attack surface: video calls, audio calls, messaging. Attackers exploit the (often) lower-security context of remote communication (less physical oversight, more familiarity). Mirage Security

4. Drift in Effectiveness: What's Changing?

From the above data and observations, we can infer several trends in *effectiveness* of AI-enabled social engineering:

- **Higher success probability per attempt**... Because personalized, convincing messages reduce user skepticism, the number of attempts needed for a successful compromise drops. (Not all metrics public, but strong qualitative indicators).
- **Increased financial/operational impact when attacks succeed**... Example: high-value wire transfers, CEO impersonation leading to large losses.
- **Reduction in attacker cost/investment**... Generative AI and automation lower the barrier to entry; attackers don't need large teams to craft personalized scams at scale.
- **Wider reach and faster turnaround**... Thousands of personalized attacks can be executed quickly, exploiting lateral movement into voice and video domains, across channels.
- **More blind spots in defense**... Traditional email-filtering or signature detection is less effective; human training is less effective when messages are highly credible; voice/deepfake impersonation is less well defended in many organizations.
- **Shift toward stealth and "trusted person" impersonation**... The attacker's persona is less obviously malicious; e.g., a voice is recognized, a familiar executive is seen in a video, reducing suspicion.
- **Layered multi-vector attacks**... Phishing + voice + video + social media reconnaissance create more complex campaigns that are harder to counter in isolation.

5. Why this Trend Matters (Implications for Organizations)

Strategic Risk

- Social engineering is no longer just "click this link" email phishing... it is *trusted-person impersonation*, cross-channel, high-impact. The risk model must evolve accordingly.
- High-value incidents (wire fraud, business compromise) can result in reputational damage, regulatory exposure, large direct losses.
- The industrialization of fraud means attackers can move from opportunistic to systematic operations.

Defense Gap

- Many organizations have invested heavily in technical controls for malware, ransomware, network intrusion, but less so for social engineering particularly in voice/video domains.
- User awareness training may lag behind: if the messages/voices seem authentic, traditional training is less effective.
- Verification processes (especially for high-value transactions) may assume "voice from a trusted exec means valid"... this assumption is now dangerously weak.

Human/Process Dimension

- The "human layer" is still the weakest link. When that link is attacked with high-fidelity social engineering (via AI), the effectiveness increases dramatically.
- Organizations must rethink transactional verification: e.g., independent confirmation, process segmentation, transaction thresholds for voice instructions, video instruction.
- Monitoring and anomaly detection: voice or video calls leading to transactions should trigger correlated checks, especially if atypical.

Regulatory / Governance

- Regulators are increasingly issuing alerts (for example, Financial Crimes Enforcement Network, US Treasury) about deepfake scams, impersonation via AI. FinCEN.gov
- Boards and audit/risk committees must consider AI-enabled social engineering as a significant emerging risk.
- Cyber insurance and liability: insurers are adjusting models for these threats, incident response plans must incorporate voice/deep-fake compromise.

6. Case-Studies & Illustrative Incidents

- A multinational firm lost **US $25 million** after a deep-fake audio/video impersonation of their CFO tricked an employee into wiring funds. SpringerLink
- Regulatory alert: Deepfakes used to mimic identities, circumvent biometric verification, and cause employees to take unauthorized actions. Department of Financial Services
- Phishing-as-a-service and AI-powered polymorphic phishing: Allot reports of such campaigns in 2024 that evade traditional controls. Allot

7. Quantitative Trend Summary (Approximate)

Metric	Approximate Trend	Notes / Sources
Increase in phishing volume	~ 50 %+ year-on-year in some studies (e.g., Allot 2024)	Allot
Increase in deepfake attacks seen by organizations	61 % organizations reported increases (2023-24)	Cobalt
Deepfake/voice-impostor growth	Deepfake vishing +1,600 % in Q1 2025 vs end-2024	Right-Hand Cybersecurity
Forecast losses due to generative-AI fraud	US$40 billion (by 2027 in US)	appgate.com

While public data on “success rate” (i.e., percent of attempts that result in compromise) is scarce, all indicators point to increasing *effectiveness* because of improved targeting, realism, multichannel attack vectors, and higher value compromises.

8. Future Outlook & Predictions (Next 3–5 Years)

- The “arms race” between defenders and attackers will move into **AI vs AI**: attackers using generative-AI to craft social engineering, defenders deploying AI/detection to identify synthetic content.
- Audio and video deepfakes will become more common in corporate fraud (video calls with executives, voice-over instructions for transfers)... defense must adapt accordingly.
- Automation of social engineering campaigns will scale massively: low-cost, high-volume, cross-language, multichannel (email, chat, voice, video).
- Process redesign and human-verification workflows will become critical: e.g., requiring secondary independent confirmation for requests made via voice or video, especially if financial or sensitive.
- Regulations, disclosure requirements, and insurance models will increasingly mandate “AI-enabled social engineering” as a separate risk category.
- Organizations that do *not* adapt (in process, human training, verification, detection technology) will see disproportionate losses. The attacker cost-barrier will flatten, making small organizations attractive targets as well.
- As generative-AI becomes ever more accessible, awareness fatigue may set in... so organizational culture, simulation-exercises, process resilience become more important than ever.

9. Recommendations for Organizations and Practitioners

1. **Elevate social engineering (including AI-enabled) to Board-level risk discussion.** Treat impersonation, voice/video deepfakes, multichannel scams as core aspects of cyber risk.
2. **Update verification processes**: For any transaction or request made via voice/video from "trusted persons", require secondary independent confirmation (preferably via a different channel).
3. **User awareness training -> advanced simulation**: Introduce training that reflects AI-generated content (voice, video, text), not just "look for typos and suspect links".
4. **Deploy detection technologies**: Utilize AI/ML tools for detecting synthetic media, abnormal voice traffic, impersonation patterns, and integrate with SOC workflows.
5. **Monitor metrics and incidents**: Record and trend data on attempted voice/value transfers, deepfake flagged contacts, etc. Build incident response playbooks for AI-enabled social engineering.
6. **Strengthen segmentation and least-privilege**: Reduce "single point victim" risk by limiting any one individual's ability to transfer large funds or to bypass controls based solely on an instruction.
7. **Collaborate with insurers, regulators and industry peers**: Share intelligence on emerging social engineering campaigns, especially those using generative AI.
8. **Consider supplier/distribution chain risk**: Attackers may impersonate not just internal execs but external partners or vendors using deepfake voice/video to exploit trust chains.

10. Conclusion

Social engineering has evolved from relatively low-sophistication, human-driven phishing into a highly automated, high-scale, AI-augmented threat vector. The convergence of generative AI (text/voice/video), massive personal data exposure, remote working modalities, and the attacker move-toward automation means that the *effectiveness* of social engineering has increased markedly... both in terms of volume and impact.

From an industry-analyst perspective with decades of witnessing shift after shift in cyber-threats, what is different now is the *speed* and *reach* of change. Where earlier defenders could adapt over months or years, today attackers can adopt AI-driven social engineering techniques almost in real time. Organizations must accelerate their own evolution in process, human-factor readiness, and detection capabilities or risk being caught unprepared.

References

1. Allot Communications. “Decoding Deception – The Rise of AI-Driven Phishing Threats” (2024) – noted 54 % increase in phishing attacks and AI-powered polymorphic phishing campaigns. Allot
2. Yale University Cybersecurity Awareness Tip (April 2024). “Personalized phishing attacks – AI bots assist criminals … Deepfake creation: AI bots can create realistic audios and videos…” Yale Cybersecurity
3. CrowdStrike. “AI-Powered Social Engineering Attacks” (2024) – how AI advances social engineering, automates phone calls, voice cloning. CrowdStrike
4. Springer article “Generative artificial intelligence in social engineering and phishing” (2024) – voice phishing (vishing) and deepfake attacks rising as attackers use generative AI. SpringerLink
5. MDPI. “Deepfake-Driven Social Engineering: Threats, Detection …” – how artificial media produced by AI can impersonate authoritative personas, reshape threat landscape. MDPI
6. World Economic Forum / article on AI agents in cybersecurity (2024) – AI-based attacks are automated, adaptive and tailored to their targets. World Economic Forum
7. Industry letter (NY DFS) “Cybersecurity Risks Arising from …” (October 16 2024) – AI-driven social engineering attacks lead to unauthorised actions such as wiring funds by mimicking voices/appearance. Department of Financial Services
8. Paubox blog “How AI is arming phishing and deepfake attacks” – “AI-generated communications nearly indistinguishable from legitimate messages.” Paubox
9. Cobalt blog “Top 40 AI Cybersecurity Statistics” – 61 % of organizations saw an increase in deepfake attacks; deepfake attacks projected to increase. Cobalt
10. Academia: Falade “Decoding the Threat Landscape: ChatGPT, FraudGPT…” (2023) – generative AI models augment social engineering attacks (phishing, impersonation). arXiv
11. Academia: Mirsky & Lee “The Creation and Detection of Deepfakes: A Survey” (2020) – early survey of deepfake creation/detection, marking evolution. arXiv