Deepfakes as a Cybersecurity Governance Problem: Legal Gaps, Institutional Risk, and Public Authority

TOMILOLA AYENI
Business Administration, University of Northampton

Abstract - Deepfakes are increasingly treated as a problem of misinformation, media ethics, or individual harm. This article argues that deepfakes should instead be understood as a systemic cybersecurity risk to public institutions. By enabling realistic impersonation without technical intrusion, deepfakes undermine authentication, disrupt public communications, and weaken institutional authority. Existing legal frameworks address some downstream harms, such as defamation or fraud, but they do not adequately account for deepfakes as a threat to secure governance. This article examines how deepfakes challenge traditional cybersecurity models, evaluates the limits of current legal responses, and argues for integrating synthetic media risks into cybersecurity law, public-sector governance, and incident response frameworks.

Keywords: Deepfakes, Cybersecurity Governance, Institutional Trust, Public Sector Cybersecurity, Synthetic Media

I. INTRODUCTION

Cybersecurity law has historically focused on protecting systems, data, and infrastructure from unauthorized access, as seen in early laws such as the Computer Security Act of 1987 (Congress, 2019; Sekara, 2020). Threat models have centered on hackers, malware, and network vulnerabilities (Cisco, 2025). Deepfakes disrupt this model by shifting the locus of attack from systems to trust. A deepfake can impersonate a public official, senior civil servant, or institutional spokesperson without breaching any technical perimeter. The resulting harm is not a data breach, but a breakdown in confidence.

That shift is not theoretical. The Ponemon–Sullivan Privacy Report (2025) finds that executives are being targeted by deepfake images or videos an average of three times, and that organizations are increasingly worried about these attacks. More than half of IT security practitioners surveyed ranked deepfakes among the most concerning uses of AI, and many organizations report being unprepared to detect or respond to them. That matters because public institutions depend on trust as a functional requirement of governance. Elections, emergency response, public health communication, and financial regulation all rely on the assumption that official messages can be verified as authentic. Deepfakes undermine that assumption by making authenticity difficult to prove and easy to fake. In practice, that means deepfakes should be treated not merely as speech or expression, but as security incidents with legal consequences, requiring new rules for attribution, liability, and institutional responsibility.

Deepfakes and the Limits of Traditional Cybersecurity Models

Most cybersecurity frameworks (NIST, HIPAA/HITRUST, etc.) are designed to protect digital assets. They assume that threats involve unauthorized access to systems or information (NIST, 2025). Deepfakes operate differently. They exploit publicly available data, such as speeches, interviews, and social media content, to construct false representations that appear authentic. The attack vector is perception rather than infrastructure.

This places deepfakes closer to social engineering than to hacking (Pedersen et al., 2025; Jampani, 2025). However, unlike phishing emails or spoofed phone calls, deepfakes can carry audiovisual credibility. A synthetic video of a minister announcing a policy change, or a fabricated audio clip of a central bank official commenting on interest rates, can have immediate real-world consequences. Markets may react, public behavior may change, and institutional legitimacy may be questioned before verification mechanisms can respond.

From a cybersecurity standpoint, this represents a failure of authentication. The audience cannot easily verify whether the speaker is genuine. Yet most legal frameworks do not treat impersonation through synthetic media as a security failure unless it results in a clearly defined offense, such as fraud.

Legal Responses: Fragmentation and Reactive Enforcement

Current legal responses to deepfakes are fragmented. Different areas of law address different harms, but none comprehensively frame deepfakes as a threat to institutional security. Criminal fraud statutes may apply when deepfakes are used to obtain money or property (Commonwealth of Pennsylvania, 2025). Defamation law may apply when reputational harm occurs (George, 2024). Data protection laws may apply when personal data is unlawfully processed (National Conference of State Legislatures, 2024).

However, these doctrines are largely reactive. They address harm after it has occurred and place the burden on affected individuals or institutions to seek redress. In the context of deepfakes, this approach is poorly suited to the speed and scale at which harm can spread. A false video can circulate globally within minutes, while legal remedies may take months or years.

Public law offers limited tools to address this gap. While some jurisdictions have begun to criminalize specific uses of deepfakes, such as non-consensual intimate imagery or election interference, these statutes are narrow and context-specific. They do not address the broader problem of deepfake impersonation targeting public authority itself.

Public Institutions as Targets, Not Just Victims

Deepfakes increasingly target institutions rather than individuals. When a deepfake impersonates a public official, the harm extends beyond personal reputation. It undermines the institution's credibility and authority as a source of reliable information. This has implications for administrative law, public accountability, and democratic legitimacy, all of which depend on the public's ability to identify authentic institutional speech.

Recent cybersecurity research confirms that this shift is structural rather than incidental. Analysis by Carnegie Mellon University's Software Engineering Institute shows that while early deepfake attacks focused on public figures, government bodies, healthcare institutions, and other public-facing organizations are now prime targets. These attacks rarely involve breaching internal systems. Instead, they exploit the communicative role of institutions by fabricating audio, video, or text that appears to originate from an official source. From a legal perspective, this matters because the attack surface is no longer technical infrastructure but institutional trust itself, a resource that existing cybersecurity law largely assumes rather than protects (Walsh, 2024).

The consequences can be immediate and severe. If a deepfake video appears to show a public health authority issuing false guidance during a crisis, the resulting confusion may cause real-world harm even after the content is disproven (Romanishyn, Malytska, and Goncharuk, 2025). Yet the legal system struggles to respond. Attribution is difficult, perpetrators may operate across jurisdictions, and many regulatory regimes still conceptualize cybersecurity incidents as breaches of data or systems rather than attacks on the authenticity of official communications. As the SEI research demonstrates, detection is technically possible but increasingly complex, requiring specialized tools and human–machine workflows that many public institutions do not currently deploy (Walsh, 2024).

From a governance standpoint, this exposes a regulatory blind spot. Public-sector cybersecurity strategies tend to focus inward, emphasizing network security, access controls, and data protection (Borky and Bradley, 2019). Far less attention is paid to safeguarding the integrity of outward-facing communications, despite their centrality to public function. As deepfake generation methods become more accessible and sophisticated, this gap raises a fundamental legal question: whether the state's duty to secure its operations should extend to protecting the authenticity of its own voice. Until cybersecurity law adapts to this reality, public institutions will remain not just victims of deepfakes but deliberate and strategically vulnerable targets.

Deepfakes and Election Security

Election integrity has become a central concern in the regulation of deepfakes. State legislatures increasingly recognize that synthetic media can distort political discourse and undermine public confidence in electoral outcomes, particularly as generative tools become cheaper and more accessible. Much of this legislation focuses on campaign conduct, disclaimers, and deceptive political advertising. What it often overlooks, however, is institutional impersonation. A deepfake that falsely depicts an election authority announcing polling changes or delayed results does not neatly fit traditional election law categories, yet it can produce immediate and widespread confusion.

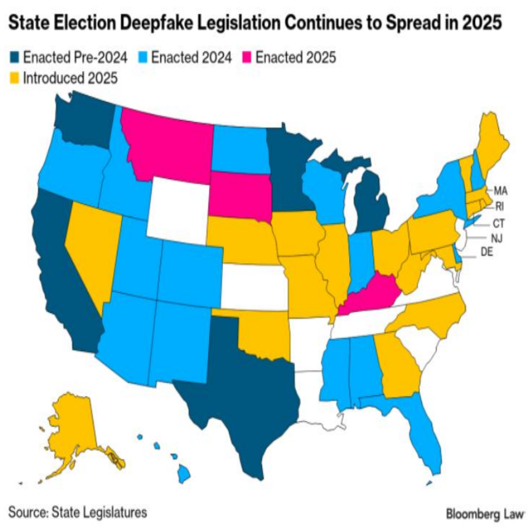
State Election Deepfake Legislation Continues to Spread in 2025
Enacted Pre-2024
Enacted 2024
Enacted 2025
Introduced 2025
MA
RI
CT
NJ
DE
Source: State Legislatures
Bloomberg Law

Figure 1: State Election Deepfake Legislation Continues to Spread in 2025 (Yuille, 2025)

As the expanding legislative map illustrates, regulation of election-related deepfakes has accelerated rapidly, with a growing number of states enacting or proposing targeted statutes. This patchwork approach reflects urgency, but it also exposes structural gaps. Many laws regulate the use of deepfakes by private actors while leaving unclear how public institutions themselves should detect, authenticate, and respond to impersonation. Courts have begun to scrutinize these statutes under the First Amendment, particularly where criminal liability or broad publication bans are involved, further complicating enforcement and leaving election officials with uncertain legal tools during fast-moving incidents.

From a cybersecurity perspective, deepfake election incidents should be treated as attacks on electoral infrastructure, even when no voting machines or databases are compromised. Election systems rely on trusted communications as much as technical integrity. A fabricated announcement attributed to an election authority can disrupt turnout, erode legitimacy, and strain emergency response mechanisms. Integrating deepfake detection, authentication of official communications, and rapid legal response into election cybersecurity planning would better align law with the realities of modern threat models. Without this shift, election law risks addressing the symptoms of synthetic media misuse while leaving the underlying institutional vulnerabilities intact.

Toward a Cybersecurity Governance Approach

Addressing deepfakes effectively requires a shift in legal framing. Rather than treating deepfakes solely as expressive content, lawmakers and regulators should recognize them as tools that can undermine secure governance. This does not require banning the technology. Instead, it calls for institutional responsibility and procedural safeguards.

Public institutions should be encouraged, or required, to adopt authentication measures for official communications. This may include verified channels, cryptographic signatures, or public education campaigns explaining how to identify genuine messages. Cybersecurity incident response plans should explicitly include deepfake scenarios and clearly define lines of authority for verification and correction.

From a legal perspective, this also raises questions about standards of care. If a public institution fails to implement reasonable safeguards against impersonation and foreseeable harm occurs, the institution may be liable. While courts have not yet clearly articulated such duties, the logic is consistent with existing principles in public-sector risk management.

II. CONCLUSION

Deepfakes expose a fundamental weakness in existing cybersecurity and legal frameworks: their reliance on technical boundaries rather than social trust. By enabling realistic impersonation without system intrusion, deepfakes challenge how public institutions establish authority and communicate with the public.

Treating deepfakes as a cybersecurity governance problem offers a more coherent response than fragmented, harm-specific regulation. It shifts attention toward prevention, institutional preparedness, and trust preservation. As synthetic media becomes more sophisticated, the law's task is not to chase every new technique, but to reinforce the conditions under which public authority remains credible.

Deepfakes do not break the law so much as they reveal where the law has not yet caught up with how power, technology, and trust now interact.

REFERENCES

[1] Borky, J.M. and Bradley, T.H. (2019). Protecting Information with Cybersecurity. *Effective Model-Based Systems Engineering*, [online] 1(1), pp.345–404. https://doi.org/10.1007/978-3-319-95669-5_10.

[2] Cisco (2025). *What Is Threat Modeling?* [online] Cisco. Available at: https://www.cisco.com/site/us/en/learn/topics/security/what-is-threat-modeling.html.

[3] Commonwealth of Pennsylvania (2025). *Gov. Shapiro Signs New Digital Forgery Law*. [online] Available at: https://www.pa.gov/governor/newsroom/2025-press-releases/gov--shapiro-signs-new-digital-forgery-law.

[4] Congress (2019). *H.R.145 - 100th Congress (1987-1988): Computer Security Act of 1987*. [online] Congress.gov. Available at: https://www.congress.gov/bill/100th-congress/house-bill/145.

[5] George, A. (2024). *Defamation in the Time of Deepfakes*. [online] Social Science Research Network. https://doi.org/10.2139/ssrn.4719803.

[6] Jampani, S.K. (2025). Social Engineering 2.0 Deepfake and Deep Learning-based Cyber-attacks (Phishing). *International Journal For Multidisciplinary Research*, 7(1). https://doi.org/10.36948/ijfmr.2025.v07i01.35527.

[7] National Conference of State Legislatures (2024). *Deceptive Audio or Visual Media (‘Deepfakes’) 2024 Legislation*. [online] www.ncsl.org. Available at: https://www.ncsl.org/technology-and-communication/deceptive-audio-or-visual-media-deepfakes-2024-legislation.

[8] NIST (2025). *Cybersecurity Framework*. [online] National Institute of Standards and Technology. Available at: https://www.nist.gov/cyberframework.

[9] Pedersen, K.T., Pepke, L., Stærmose, T., Papaioannou, M., Choudhary, G. and Dragoni, N. (2025). Deepfake-Driven Social Engineering: Threats, Detection Techniques, and Defensive Strategies in Corporate Environments. *Journal of Cybersecurity and Privacy*, [online] 5(2), p.18. https://doi.org/10.3390/jcp5020018.

[10] Ponemon Institute Research (2025). *Deepfake Deception: How AI Harms the Fortunes and Reputations of Executives and Corporations | Ponemon-Sullivan Privacy Report*. [online] Ponemonsullivanreport. Available at: https://ponemonsullivanreport.com/2025/04/deepfake-deception-how-ai-harms-the-fortunes-and-reputations-of-executives-and-corporations/.

[11] Romanishyn, A., Malytska, O. and Goncharuk, V. (2025). AI-driven disinformation: policy recommendations for democratic resilience. *Frontiers in Artificial Intelligence*, 8. https://doi.org/10.3389/frai.2025.1569115.

[12] Sekara, H. (2020). *A Look at the Computer Security Act of 1987 | Tripwire*. [online] www.tripwire.com. Available at: https://www.tripwire.com/state-of-security/computer-security-act-of-1987.

[13] Walsh, M. (2024). A Framework for Detection in an Era of Rising Deepfakes. *Cybersecurity Engineering*. https://doi.org/10.58012/e5sh-hp94.

[14] Yuille, T. (2025). *ANALYSIS: Deepfake Election Laws Spread Amid Court Challenges*. [online] @BLaw. Available at: https://news.bloomberglaw.com/bloomberg-law-analysis/analysis-deepfake-election-laws-spread-amid-court-challenges-12.