FROM DISINFORMATION TO SYNTHETIC REALITY: DEEPFAKES AS INFRASTRUCTURAL POWER IN DIGITAL POLITICAL COMMUNICATION

Anahit Menemshyan

PhD, Associate Professor

Email: a.menemshyan@ysu.am,

ORCID: 0009-0002-9403-4975

Varduhi Petrosyan

PhD, Associate Professor

Email: varduhipetrosyan@ysu.am

ORCID: 0009-0006-5182-2890

Abstract

The rapid development of generative artificial intelligence is transforming the production, circulation, and evidentiary status of political communication. Deepfakes challenge established assumptions about the authenticity of audiovisual evidence. This article examines deepfakes not merely as disinformation, but as synthetic impersonation that can fabricate the apparent speech, voice, behavior, or participation of a political actor. The study shifts the analytical focus from whether a media item is fake to how political authenticity is manipulated.

Drawing on a qualitative comparative case study of three publicly documented deepfakes circulated before Armenia's 2026 parliamentary elections, the article identifies three mechanisms of synthetic impersonation: synthetic attribution, synthetic recontextualization, and synthetic vocal attribution. The analysis demonstrates that the primary object of manipulation is not simply information, but the perceived relationship between a political actor and an audiovisual representation of their speech or action. This creates a broader challenge to political communication: deepfakes can make fabricated evidence appear authentic while contributing to the possibility that authentic recordings may be dismissed as artificial.

The article employs infrastructural power as an interpretive framework for examining how synthetic impersonation extends beyond the fabricated media artifact to digital environments in which content is circulated, verified, and contested. The study does not claim causal effects of algorithms or platforms, but situates deepfakes within a wider platform-mediated communicative infrastructure. The findings highlight implications for political authenticity, evidentiary credibility, electoral integrity, and public trust, underscoring the importance of transparency, independent fact-checking, platform cooperation, and media literacy for democratic communication.

Keywords: deepfakes; synthetic media; synthetic impersonation; political communication; political authenticity; audiovisual credibility; infrastructural power; electoral integrity

Introduction

The rapid advancement of artificial intelligence (AI), particularly generative AI, has expanded the technological capabilities of digital communication and reconfigured the conditions under which political messages are produced, circulated, and interpreted. Systems capable of generating or substantially modifying text, images, audio, and video challenge established assumptions regarding the evidentiary value of audiovisual media. This shift is especially consequential for political

communication, where message credibility often hinges not only on the content itself, but also on who is depicted as the speaker and the extent to which the recording appears authentic.

Deepfakes represent one of the most significant manifestations of this transformation. By synthesizing or manipulating an individual's face, voice, speech, or actions, they can generate audiovisual representations that depict events that never occurred or statements that were never made.

Existing scholarship has largely examined deepfakes as a form of disinformation, a technical detection challenge, or an ethical and regulatory problem. However, these frameworks do not fully capture another communicative issue that is particularly critical in electoral contexts: deepfakes can fabricate the apparent presence of a political actor within a communicative event.

Accordingly, this article shifts the analytical focus from the question of "whether a given item is fake" to "which mechanism of political authenticity is being manipulated." The central argument is that deepfakes should be understood as a form of synthetic impersonation; they not only generate false information about a political actor, but also fabricate their apparent speech, voice, behavior, or participation in a specific event. This alters the evidentiary conditions of political communication, as the credibility of audiovisual evidence has traditionally relied, in part, on the perceived continuity between the individual and the recorded action.

While infrastructural power is retained as a theoretical lens, the article does not frame it as a causal property directly proven by the three case studies. Rather, it helps explain why the political significance of synthetic impersonation cannot be isolated from the digital environments in which deceptive content becomes visible, re-circulating, and contested. The empirical analysis focuses on observable communicative and evidentiary mechanisms within documented cases.

The study is guided by two primary research questions: (1) how do deepfakes alter the mechanisms through which political authenticity is constructed and contested, and (2) what forms of synthetic impersonation are observable in documented cases of deepfakes during the period preceding Armenia's 2026 parliamentary elections?

Based on this framework, the study hypothesizes that political deepfakes primarily disrupt political communication by manipulating the perceived relationship between a political actor and an audiovisual representation of their speech, identity, or actions, rather than merely by introducing false information. More specifically, the study expects that documented cases of political deepfakes can be distinguished according to different mechanisms of synthetic impersonation, including synthetic attribution, synthetic recontextualization, and synthetic vocal attribution.

Theoretical Framework

From Information Disorder to Synthetic Reality

The concept of information disorder has introduced a critical distinction into the study of problematic information, moving beyond a simple true/false binary. Wardle and Derakhshan differentiate among misinformation, disinformation, and malinformation by considering the interplay of falsity, intent, and harm. Disinformation is particularly central here, as the cases under examination involve materials deliberately created or manipulated within a political context (Wardle & Derakhshan, 2017).

Digital platforms have transformed disinformation into a sociotechnical process. Social media platforms, recommendation systems, engagement mechanisms, and platform governance condition not only the speed at which information circulates, but also its visibility and contextualization (Gillespie, 2018; van Dijck et al., 2018). Consequently, the study of disinformation is increasingly shifting away from isolated pieces of content toward the infrastructures and communicative

environments through which content acquires political significance. Deepfakes intensify this shift, as the object of manipulation is no longer merely an assertion, but the audiovisual evidence that ostensibly substantiates it.

Deepfakes and the Crisis of Audiovisual Credibility

Synthetic media refers to content generated or substantially modified through artificial intelligence. Deepfakes represent a distinct subcategory of synthetic media, as they simulate the identity or presence of real individuals. Advances in generative models, including generative adversarial networks (GANs), have significantly heightened the realism of synthetic audiovisual materials (Goodfellow et al., 2014; Westerlund, 2019; Mirsky & Lee, 2021).

However, the significance of deepfakes extends beyond technical realism. Audiovisual media has traditionally possessed evidentiary weight because a recording was understood to preserve a trace of an event—the individual appeared to have spoken, acted, or been present. Deepfake manipulations disrupt this connection between representation and event. A deceptive recording can simulate not only what was said, but also the apparent fact that a specific individual said it. Consequently, the central issue becomes one of authenticity: who is depicted as the speaker, whether the depicted action actually occurred, and what status the recording can legitimately hold as evidence (Fallis, 2021).

Recent experimental research similarly demonstrates that audiences evaluate political deepfakes not merely on the basis of message content, but through the perceived authenticity of its presentation (Hameleers et al., 2025). This distinguishes political deepfakes from conventional textual disinformation: manipulation can operate simultaneously at the level of both the assertion and the evidentiary representation.

Infrastructural Power and Algorithmic Mediation

The political significance of deepfakes depends heavily on the communicative infrastructures through which they circulate. Research on platforms and infrastructures demonstrates that digital architectures condition how information is produced, distributed, and rendered visible (Plantin et al., 2018; Gillespie, 2018). While platform power denotes a platform's direct capacity to rank, moderate, remove, or otherwise structure content, infrastructural power operates at a foundational level, governing the conditions under which communication becomes possible and visible.

This conceptualization helps reframe deepfakes not as isolated media artifacts, but as integrated components of a broader communicative ecosystem. Attention-driven platform logics tend to amplify novel, emotionally charged, and conflictual content, while algorithmic curation structures user exposure (Srnicek, 2017; Tufekci, 2017; van Dijck et al., 2018). The present study does not claim to empirically quantify these platform mechanisms within the Armenian cases; rather, infrastructural power serves as a meso-level interpretive framework to explain why the ramifications of synthetic impersonation extend well beyond the boundary of the original fabricated artifact.

Methods

The empirical section employs a qualitative comparative case study design. The corpus comprises three publicly documented cases of political deepfakes circulated during the period leading up to Armenia's 2026 parliamentary elections. Cases were selected based on four criteria: (a) the material incorporated AI-generated or AI-manipulated audiovisual content; (b) a real political or public figure was depicted saying or doing something that did not actually occur; (c) the item served an explicit political-communicative function; and (d) its deceptive nature was independently verified by a fact-checking organization or confirmed through direct comparison with the primary source material.

The cases are analyzed across four distinct dimensions: (1) mode of synthetic manipulation (facial, vocal, textual, or contextual alteration); (2) target of manipulation (a specific political actor

or political discourse more broadly); (3) communicative function (including character assassination, false attribution, or mobilization); and (4) impact on political authenticity (whether the alteration aims to pass false evidence off as genuine, alter the speaker's identity, or erode trust in authentic evidence). Accordingly, the analysis remains interpretive and evidence-based; it does not attempt to evaluate audience reach, voter belief, electoral impact, or platform-driven causal amplification unless such data were explicitly documented in the available record.

Results

Deepfakes in Armenian Pre-Election Communication: Three Cases

The cases presented below were selected not only for their political significance, but also because their AI-driven or digital manipulation was independently verified through fact-checking. Together, they represent distinct forms of synthetic impersonation, offering a basis for comparing the mechanisms through which political authenticity is manipulated.

False Political Speech: The Video of Nikol Pashinyan

In April 2025, a fraudulent video circulated on social media depicting Prime Minister Nikol Pashinyan using offensive language directed at Elon Musk. FIP.am confirmed that the material was generated using AI and that Pashinyan had made no such statement. Fact-checkers identified several indicators of manipulation: unnatural speech patterns, discrepancies between facial expressions and movements, synthetic voice artifacts, and a high probability score for deepfake generation returned by automated detection software. The content thus constructed a false political utterance and attributed it to a real political figure.

The primary communicative mechanism in this case is synthetic attribution. Rather than merely fabricating a claim, the manipulation constructs the apparent fact that the politician himself uttered it. Consequently, its political function is character assassination. The deceptive statement gains heightened efficacy precisely because it is presented through the simulated voice and likeness of the leader. Through the analytical framework, this case links audiovisual impersonation, false attribution, and reputational damage.

Recontextualized Identity: "Matryoshka" Deepfakes

In May 2026, CivilNetCheck reported on a series of deepfake videos circulated via the Russian-linked network known as "Matryoshka." One such video utilized footage of Sergeant Steve Arnold from the South Australia Police. The original video released by the police department highlighted his work as a mentor and contained no reference whatsoever to Armenia or Nikol Pashinyan. In the manipulated iteration, AI was used to insert mentions of Pashinyan alongside incitements to violence—elements completely absent from the primary source. Another video appropriated footage of Caroline Clarke, Regional Director for NHS London, falsely attributing to her claims about Armenia and a political call to vote against Pashinyan. No such statements existed in the original recording.

These cases demonstrate a second mechanism: the synthetic recontextualization of a real individual within genuine footage. The political message derives its apparent credibility from the fact that the underlying visual source is real, even though the audio and context have been falsified. Thus, the manipulation operates at the intersection of authentic evidence and synthetic attribution. It is not necessary to fabricate an entire audiovisual event from scratch; strategically altering an authentic individual's apparent speech is sufficient to construct a deceptive political narrative.

Synthetic Voice Attribution: The Samvel Karapetyan Case

In June 2026, FIP.am reported on an AI-generated video circulated as paid advertising on YouTube, in which a synthetic voice was attributed to political figure Samvel Karapetyan. The video

incorporated footage from his May 7, 2026 press conference but assigned him statements he had never made. Fact-checking verified that the attributed claims were entirely absent from the original event. This case highlights the growing role of voice cloning in political communication: the visual source may remain authentic, while the speaker's apparent voice and speech are synthetic.

The primary mechanism here is synthetic vocal attribution. Political impact does not require the fabrication of an entirely new visual event; rather, it leverages the recognizable image of a public figure along with the evidentiary authority traditionally associated with their voice. Furthermore, the case demonstrates that deepfakes can serve not only direct persuasion, but also ridicule and delegitimization by attributing statements to a political actor that compromise their public image or credibility.

Discussion

Comparative Discussion: From False Information to Synthetic Presence

Although distinct in their technological mechanics, these three cases reveal a shared communicative structure. In the first instance, a political statement is fabricated and directly attributed to a target. In the second, authentic footage of real individuals is recontextualized to depict them making political declarations they never actually delivered. In the third, genuine visual source material is paired with a synthetic voice to execute a deceptive vocal attribution. Across all three cases, the primary object of manipulation is not merely information, but communicative presence itself.

This distinction holds significant theoretical weight. Traditional disinformation may fabricate a claim while leaving the identity of the source intact. Deepfakes, by contrast, falsify the nexus between the source and the utterance. They address a fundamental communicative question—not merely "is this statement true?" but "did this individual actually say or do this?" Consequently, the political target undergoes synthetic impersonation, wherein identity itself is weaponized as a vehicle for disinformation.

These cases can be conceptualized as three interconnected transformations of political authenticity: synthetic vocal attribution, synthetic identity impersonation, and the synthetic recontextualization of authentic audiovisual material. Their cumulative effect is the erosion of the evidentiary distinction between a genuine recording and a fabricated representation. This does not imply that audiences uncritically accept every deepfake; rather, the underlying political challenge lies in the destabilization of the broader evidentiary environment.

This dynamic directly aligns with the "liar’s dividend" outlined by Chesney and Citron (2019). As deepfakes become a widely recognized possibility, political actors can exploit the resulting ambiguity to dismiss authentic recordings as fabrications. Thus, the technology functions as a dual-edged sword: it elevates the perceived credibility of deceptive evidence while simultaneously undermining trust in legitimate evidence. Recent research on political deepfakes similarly underscores that presentation authenticity constitutes a distinct dimension of credibility, independent of the intrinsic plausibility of a given message (Hameleers et al., 2025).

The infrastructural lens introduces a final analytical layer to this interpretation. While these cases alone do not provide sufficient empirical grounds to claim that specific recommendation algorithms or platform architectures causally drove their diffusion, they demonstrate why the political significance of synthetic impersonation cannot be divorced from digital environments characterized by rapid recirculation, cross-platform reproduction, and contested visibility. Accordingly, the proper unit of analysis is not merely the fabricated file itself, but the entire communicative circuit linking production, circulation, verification, contestation, and public interpretation.

Regulatory and Ethical Implications

The empirical cases point to a limited set of regulatory and ethical issues directly associated with political deepfakes. The first is transparency: audiences ought to be informed when political audiovisual content has been generated or substantially altered via AI. The second is responsibility: accountability cannot be attributed solely to technical producers when manipulated material is intentionally disseminated as political communication. The third is electoral integrity: the circulation of synthetic impersonations during elections can hinder citizens' capacity to evaluate political actors on the basis of their authentic statements and actions.

The European Union's AI Act provides an important reference framework for transparency by imposing disclosure obligations on certain AI-generated or manipulated content (European Parliament & Council of the European Union, 2024), while the Digital Services Act addresses systemic risks on very large online platforms (European Parliament & Council of the European Union, 2022). International frameworks, such as UNESCO's Recommendation on the Ethics of Artificial Intelligence and the OECD AI Principles, similarly emphasize transparency, accountability, human oversight, and the protection of democratic values (Organisation for Economic Co-operation and Development, 2019; UNESCO, 2021).

In the Armenian context, the critical takeaway is not a blanket prohibition of synthetic media-which could pose severe challenges to freedom of expression-but rather the development of clear disclosure, verification, and accountability mechanisms tailored to political synthetic media. Armenia's legal framework already incorporates key transparency principles governing synthetic audiovisual content, while fact-checking initiatives provide an additional layer of verification. Nonetheless, the empirical cases highlight a practical gap between formal disclosure requirements and content actually circulating across social and advertising platforms. Consequently, safeguarding electoral communication requires transparent labeling, independent fact-checking, platform cooperation, swift verification procedures, and media literacy. These measures should aim to protect the evidentiary conditions of political communication rather than suppress synthetic media as a technological category.

Unlike many international jurisdictions that have already introduced specific rules in this area, Armenia still does not have a separate, comprehensive legal framework governing the circulation of political content created with AI. However, the regulatory developments of recent years suggest that Armenia is gradually moving toward a more transparency-based approach.

Armenian media legislation already reflects the principle of transparency when it comes to synthetic audiovisual content, while fact-checking initiatives provide an additional layer of protection against misleading information. In particular, amendments to the RA Law on Audiovisual Media introduced a requirement to label synthetic content. Under Article 9.1, content that is created or modified using AI must be clearly and visibly or audibly labeled, making it possible to distinguish AI-generated content from other forms of digital editing or manipulation (Republic of Armenia, 2020, Art. 9.1).

An important aspect of this approach is that Armenia has not chosen to ban synthetic media altogether. Instead, the focus is on ensuring that people know when the content they are seeing or hearing has been created or altered by AI. This approach recognizes that synthetic media is not necessarily harmful in itself; the main concern is whether people are given enough information to understand what they are consuming.

At the same time, the current rules have an important limitation: their scope is relatively narrow. They mainly apply to audiovisual media services, while a large amount of political content is shared through social media and other digital platforms. As a result, there is still a practical gap between the

formal requirement to label AI-generated content and the political content that is actually circulating on social media and advertising platforms.

For this reason, protecting electoral communication effectively requires more than just labeling requirements. It also calls for independent fact-checking, cooperation with digital platforms, fast verification procedures, and stronger media literacy. The goal of these measures should not be to restrict synthetic media simply because it is AI-generated. Rather, they should ensure that voters can understand, evaluate, and make informed decisions about the political information they encounter.

Conclusion

The analysis demonstrates that political deepfakes should not be understood merely as a technologically advanced form of disinformation. Their defining communicative property is the ability to fabricate the apparent presence of a political actor within an audiovisual event. The three Armenian cases examined here reveal distinct manifestations of synthetic impersonation: synthetic attribution, synthetic recontextualization of authentic footage, and synthetic vocal attribution.

The resulting challenge is, therefore, a crisis of political authenticity. Deepfakes can attribute false statements to real actors by presenting them as authentic while simultaneously creating conditions under which genuine recordings can be contested as artificial. Consequently, the significance of synthetic media extends beyond individual acts of deception to touch upon the broader evidentiary conditions of political communication. It is precisely for this reason that deepfakes matter for electoral integrity and public trust.

The study also clarifies the appropriate role of infrastructural analysis. Deepfakes operate within platformied communicative environments whose visibility, circulation, and contestation are shaped by digital architectures. However, the presented cases do not justify causal claims regarding algorithmic amplification or electoral impact. Thus, infrastructural power is most accurately maintained as an interpretive framework rather than presented as an empirically proven outcome of the cases.

The broader implication is that the response to political deepfakes must move beyond technical detection alone. Safeguarding democratic communication requires preserving the link between political actors and their authentic speech, actions, and audiovisual evidence. In this sense, the central challenge posed by deepfakes is not simply that false information can appear real, but that political presence itself can be synthetically produced.

References

Chesney, R. M., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review, 107*(6), 1753–1820. https://doi.org/10.15779/Z38RV0D15J

CivilNetCheck. (2026, May 30). *Russia's "Matryoshka" is spreading deepfakes that contain calls to kill Pashinyan*. CivilNet. https://civilnet.am/en/news/1012122

European Parliament & Council of the European Union. (2022). *Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a single market for digital services and amending Directive 2000/31/EC (Digital Services Act).* Official Journal of the European Union, L 277, 1–102. https://eur-lex.europa.eu/eli/reg/2022/2065/oj/eng

European Parliament & Council of the European Union. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act).* Official Journal of the European Union, L 2024/1689. https://eur-lex.europa.eu/eli/reg/2024/1689

Fallis, D. (2021). The epistemology of deepfakes. *Synthese, 198*(1), 1–23.

FIP.am. (2025, April 25). *The video about Pashinyan on X is a blatant deepfake*. https://fip.am/en/41956

FIP.am. (2026). *Samvel Karapetyan's voice was faked using AI: The video is being promoted through paid advertising*. https://fip.am/49338

Gillespie, T. (2018). *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems, 27*, 2672–2680.

Hameleers, M., van der Meer, T. G. L., & Dobber, T. (2025). How far can political deepfakes credibly deviate from reality? Responses to political deepfakes with varying degrees of deception. *International Journal of Communication, 19*, 180–198. https://ijoc.org/index.php/ijoc/article/view/22704

Mirsky, Y., & Lee, W. (2021). The creation and detection of deepfakes: A survey. *ACM Computing Surveys, 54*(1), Article 7. https://doi.org/10.1145/3425780

Organisation for Economic Co-operation and Development. (2019). *Recommendation of the Council on Artificial Intelligence* (OECD/LEGAL/0449). https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

Plantin, J.-C., Lagoze, C., Edwards, P. N., & Sandvig, C. (2018). Infrastructure studies meet platform studies in the age of Google and Facebook. *New Media & Society, 20*(1), 293–310. https://doi.org/10.1177/1461444816661553

Republic of Armenia. (2020). *Law of the Republic of Armenia on Audiovisual Media* (HO-395-N). https://www.arlis.am/en/acts/196476

Srnicek, N. (2017). *Platform capitalism*. Polity Press.

Tufekci, Z. (2017). *Twitter and tear gas: The power and fragility of networked protest*. Yale University Press.

UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. https://www.unesco.org/en/legal-affairs/recommendation-ethics-artificial-intelligence

van Dijck, J., Poell, T., & de Waal, M. (2018). *The platform society: Public values in a connective world*. Oxford University Press. https://doi.org/10.1093/oso/9780190889760.001.0001

Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policy making*. Council of Europe. https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html

Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review, 9*(11), 39–52. https://doi.org/10.22215/timreview/1282