

NB: all names, plotlines, and referenced publications, resources, archive documents and artefacts in the examples are fictional.

THE USE OF DEEPPAKE TECHNOLOGIES FOR INFLUENCING ELECTORAL PROCESSES AND THEIR IMPACT ON NATIONAL SECURITY

Assoc. Prof. Dr Gergana Varbanova

Nikola Vaptsarov Naval Academy

Abstract. The study examines the technological phenomenon of deepfake as a tool for influencing electoral processes and its impact on national security through the formation and manipulation of attitudes of individuals and social groups toward the electoral process. In the context of increasing digitalization of society and the growing influence of social media, synthetically generated content is gaining significant potential to shape public opinion, particularly during sensitive periods such as election campaigns. The article analyzes the technological characteristics of deepfake technologies and their primary applications in undermining trust in electoral processes and in the actors involved, as well as the specific mechanisms through which voter attitudes are influenced. Particular attention is given to the risks to national security, including the erosion of public trust, the potential for external interference through information operations, and the capacity for destabilization of the social environment. The study also examines the challenges associated with the legal regulation of deepfake technologies, with reference to relevant European initiatives and the existing national legal framework. It concludes by arguing for the necessity of a comprehensive approach incorporating legal, technological, and institutional measures aimed at limiting the misuse of synthetic content and ensuring the security and integrity of electoral processes.

Keywords: deepfake, electronic evidence, electoral processes, national security, synthetic content

Introduction

With the process of digital transformation of society, social media are displacing the traditional channels of information to which people have been accustomed. Information is no longer primarily obtained from traditional media such as television, radio, and newspapers, but from social media platforms such as Facebook, Instagram, and TikTok. Information disseminated through social media spreads significantly faster and reaches a wide audience, thereby increasing its influence on the individual and, consequently, on public attitudes. Technologies have thus become a key means for shaping political behavior and influencing decision-making processes, including the formation of personal choice and voting for a particular candidate or political party (Rajput 2026).

The development of artificial intelligence systems has led to the creation of generative models capable of producing synthetic content—deepfakes. This technology has rapidly evolved from

entertaining clips with satirical content into a qualitatively new type of threat, enabling the creation of content that is difficult to distinguish from authentic material. Unlike traditional disinformation, this phenomenon does not merely involve the distortion of facts, but the creation of an entirely new, synthetic reality that cannot be distinguished from authentic content.

In the context of electoral processes, such content becomes a tool for influencing public attitudes and undermining trust in democratic mechanisms.

Materials and Methods

The study aims to analyze deepfake technologies as a tool for influencing electoral processes. The focus is placed on the existing legal framework within the European Union and the associated risks to national security. The study seeks to identify deficiencies in the supranational legal framework and to substantiate the need for the development of national law in the field of electoral processes through the introduction of clear rules governing the use of synthetic content.

The present study is based on a normative (dogmatic) legal method, through which the existing legal framework at the European level, as well as relevant legislative developments in the United States, is examined. A comparative legal method is applied in order to assess and contrast different regulatory approaches to deepfake technologies. The research further adopts an interdisciplinary approach, integrating legal and technological perspectives to analyze the impact of synthetic content on electoral processes and national security.

The study aims to analyze deepfake technologies as an instrument for affecting national security through influence on individuals or groups within society and through the compromise of electoral processes. It places particular emphasis on the legal framework as a critical element in identifying, qualifying, and limiting the risks associated with synthetic content. The focus of the article is a comparative legal analysis of the regulatory approach in the United States and the supranational legislation of the European Union, including the Digital Services Act, the Artificial Intelligence Act, Section 20010 of the California Elections Code, and the limitations arising from Section 230 of the Communications Decency Act.

The analysis further examines the implications of this phenomenon for trust in electoral processes and the cognitive security of voters, as well as the challenges related to evidentiary standards, technological verification, and the so-called “non-existing deepfake”. On this basis, it substantiates the need for the development of national electoral law through the introduction of clear rules on the use, labeling, evidentiary assessment, and restriction of synthetic content in an electoral context. The article does not aim to examine the technological aspects of deepfakes in detail, as it is focused on the current regulatory framework and the challenges faced by legislative bodies in different jurisdictions.

Deepfake Technologies, Electoral Processes and National Security

Deepfakes represent synthetic audio-visual content through which words and actions are attributed to real individuals that they have never performed. In the context of electoral

processes, such synthetic content constitutes a key instrument for influencing public attitudes and undermining trust in democratic values and mechanisms. The creation of compromising synthetic content may damage not only the reputations of specific candidates, but also distort voters' perceptions and influence their electoral behavior, thereby directly affecting election outcomes. Furthermore, the dissemination of disinformation regarding the electoral process itself may influence individuals' decisions to exercise their right to vote, reduce voter turnout, and distort election results. In a broader sense, this poses risks to national security through the potential destabilization of the electoral system and the democratic foundations upon which it is built.

Synthetic content exerts a direct influence on individuals, while its dissemination through social media has been recognized as an effective means of manipulating electoral processes through disinformation or the shaping of specific attitudes. As a result, individuals are no longer able to fully trust what they see and hear (Dan 2025). This significantly complicates both the process of making informed decisions and the ability to identify synthetic content as such, as well as the formation of a critical attitude toward it.

The existing supranational legal framework at the European Union level creates more challenges than it resolves in addressing the growing volume of generative synthetic content. The main difficulties concern not only the establishment of authenticity, authorship, and the role of expert analysis as a means of distinguishing authentic from synthetic content, but also the broader crisis of trust generated by such content. Identifying hybrid threats, U.S. legislation treats the electoral process as part of critical infrastructure, as its security is directly linked to the functioning of democratic foundations and public trust in electoral outcomes. Threats are no longer limited to traditional cyberattacks targeting voter registries or technical infrastructure, but extend to social media and the information flows they generate, aimed at influencing the will of voters. In this context, deepfake technologies are identified as a new type of risk, directed not at the system itself but at the perception of its legitimacy.

Unlike conventional attacks, which target end devices as technical elements of the electoral process or involve direct intrusions into information systems processing election results, deepfake technologies operate at the cognitive level—affecting trust, perception, and the informed choice of voters. This raises the question of expanding the concept of electoral process protection to include the information environment as an element of the critical infrastructure.

New forms of influence through synthetic content cannot be effectively countered solely by technical measures. A comprehensive approach is required, combining legal and institutional mechanisms aimed at limiting manipulation in the information environment and ensuring trust in the electoral process as a fundamental element of national security.

Ewa Michałkiewicz-Kądziela identifies five structural reasons why deepfakes constitute a qualitatively more serious problem than “traditional” forms of disinformation: the increasing inability of individuals to distinguish manipulated from authentic content; the exceptionally broad potential for dissemination; the accessibility and low cost of the technology; the capacity of AI to profile and personalize influence on specific target groups; and the lack of reliable technical tools for the automatic detection and labeling of synthetic content (Michałkiewicz-Kądziela 2024). It is precisely the latter that reveals a fundamental

contradiction faced by legal regulation: it seeks to govern a phenomenon for whose identification neither society nor law enforcement authorities possess reliable instruments.

Comparative Legal Analysis: California Elections Code, Digital Services Act and AI Act

By identifying synthetic content as a threat, the California Elections Code (CEC) directs its efforts toward mitigating its impact through the adoption of legislative measures (California Elections Code § 20010). Section 20010 of the CEC delineates three forms of prohibited conduct involving deepfake manipulation and the dissemination of deceptive audio-visual content concerning a candidate within an ongoing electoral process. The first form encompasses the discrediting of a candidate through content aimed at damaging their reputation. The provision requires that such content be created with intent, specifically “to injure the candidate’s reputation.” The second form concerns the direct manipulation of electoral attitudes by political actors themselves. Deepfake content thus functions not only as a means of attacking an opponent, but also as a tool for mobilizing one’s own electorate through the construction of persuasive narratives, including the use of images of foreign political figures. The third form may also be understood as part of a process of hybrid warfare, insofar as it involves foreign interference through coordinated disinformation campaigns. In this context, deepfake technologies are used as an instrument for influencing public attitudes from outside the domestic sphere, thereby extending the issue beyond the electoral process and into the domain of national security, including the conceptualization of cyberspace as a fifth domain of warfare. In such cases, Section 20010 proves to be practically inapplicable, as it does not provide for extraterritorial effect, does not regulate the liability of foreign actors, and does not address the immunity of platforms under Section 230 of the Communications Decency Act (CDA), through which such content is disseminated to a mass audience.

A significant structural problem also lies in the fact that the primary channels of dissemination, social media platforms and paid political advertising, remain outside the effective scope of regulation. The preservation of immunity under Section 230, together with the established exceptions, effectively limits the possibility of holding platforms liable, even though they serve as the main conduits through which such content reaches a broad audience (Citron and Wittes 2017, pp. 401–456).

Despite the legislatively balanced approach of Section 20010 of the California Elections Code (CEC), its effectiveness is constrained by the federal immunity provided under Section 230 CDA, which shields internet platforms from liability for content generated by their users. Individual states have sought legislative solutions in attempts to sanction the intentional dissemination of deepfakes for disinformation purposes; however, state-level regulation encounters a “federal barrier” that prevents affected candidates from pursuing legal action against social media platforms, which constitute the primary channels for the mass dissemination of deepfake content. This renders state legislation an instrument of limited reach, incapable of effectively countering large-scale disinformation campaigns and foreign interference as a form of hybrid warfare (Nikolov and Nikolov 2024, pp. 385–390), since intermediary platforms are legally protected from civil liability and state law cannot derogate from Section 230 of the CDA.

In conclusion, a significant deficiency of the legislative approach lies in the limited temporal scope of Section 20010 of the California Elections Code, which restricts the availability of legal protection to the period of 60 days preceding the scheduled election day. This legislative design is grounded in the assumption that electoral campaigns constitute short-term, temporally confined activities; however, it fails to account for the dynamics of the digital environment and the continuous flow of information, which is often strategically deployed for disinformation or reputational harm months, and in some cases even years, prior to the electoral process itself. Consequently, the 60-day limitation renders Section 20010 a structurally insufficient instrument for addressing such phenomena through judicial intervention.

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 – DSA (Regulation (EU) 2022/2065), is not merely an administrative instrument but a functioning mechanism for the protection of cognitive security (Fathaigh et al. 2025). Synthetic content undermines public trust and, consequently, calls into question the authenticity of any content, even when it is genuine. In this context, the risk-based approach (Articles 34 and 35) obliges platforms to assess not only individual pieces of content, but also the potential of their interfaces to function as incubators of large-scale visual disinformation, affecting a broader range of social relations and even democratic processes.

The effectiveness of measures against deepfakes is further strengthened through the reconfiguration of “notice and action” mechanisms (Article 16). Given that the time window for responding to a deepfake attack against a political figure is critically short, the DSA imposes requirements of speed and accountability that are difficult to achieve within traditional judicial systems. In this respect, any natural or legal person may notify a platform and request the removal of specific content. The obligation to provide reasons for decisions (Article 17) prevents arbitrary censorship, while at the same time ensuring that manipulated content is either clearly labelled or removed.

The most significant legislative achievement of the Regulation lies in its approach to the governance of algorithmic amplification (Articles 27 and 38). Deepfake content is often designed to trigger strong emotional reactions, which engagement-driven algorithms tend to prioritize through recommender systems. By requiring transparency of these systems and the provision of alternatives not based on profiling, namely that at least one of the options offered to users must not rely on profiling, the DSA seeks to limit the spread of so-called “viral” content, which is inherently the objective of disinformation-driven synthetic campaigns. In this way, the Regulation treats deepfakes not merely as isolated media content, but as a systemic technological risk, the management of which requires transparency in AI systems and strict accountability of content-distributing platforms.

Despite legislative efforts to keep pace with the dynamic development of technologies, a fundamental challenge lies in the identification and proof that a given piece of content is synthetic, insofar as this process is entirely dependent on the adequacy of technological verification (Varbanova 2025, pp. 323–327). In the context of Section 20010 of the CEC, the requirement of “clear and convincing evidence” is difficult to meet in the absence of reliable, technological, and standardized tools for distinguishing authentic from synthetic content, thereby creating a real risk of inefficiency and limiting the effectiveness of judicial control.

Similarly, within the framework of the DSA, the lack of precise tools for automated detection calls into question the operational effectiveness of “notice and action” mechanisms (Article 16) and systemic risk assessments (Article 34). While the legislative approach emphasizes the classification of deepfake technologies as a systemic risk, the absence of technological standards for proof shifts the burden onto subjective human judgment and expert intervention, turning the fight against disinformation into a continuous race between the capabilities of artificial intelligence systems used to generate synthetic content and those used to identify such content.

The DSA is primarily directed at platforms, the dissemination of content, and the management of systemic risks in the digital environment, including the manipulation of public opinion through content amplification. By contrast, the Artificial Intelligence Act focuses on synthetic content itself as a legal phenomenon and provides a legal definition of it. At the same time, it sets out minimum requirements for disclosure and labeling, as well as the circumstances in which content must be recognizable as artificially generated or as authentic content manipulated through artificial intelligence systems.

In the context of the use of platforms to influence electoral processes, Ruling No. 32/2024 of the Constitutional Court of Romania (Constitutional Court of Romania 2024) is of particular interest, as it identifies algorithmic amplification of content as a means of manipulating voters and directly interfering with the electoral process. In its reasoning, the Court held that the State must counter organized disinformation campaigns capable of affecting the integrity of electoral processes, expressly referring to the Interpretative Declaration to the Code of Good Practice in Electoral Matters on Digital Technologies and Artificial Intelligence, adopted by the Venice Commission (para. 10). The Court further found that, through algorithmic amplification and the influence of platforms, the free formation and expression of voters’ will had been impaired by a disinformation campaign in which one of the candidates used online amplification of his campaign by circumventing national electoral legislation and abusing the algorithms of social media platforms (para. 14). The Court reasoned that the candidate’s violations had created an evident inequality between him and the other participants in the electoral process, since the significant exposure of one candidate reduced the online visibility of the others (para. 16). Although the candidate did not use synthetic content, he relied on digital technologies and the algorithmic mechanisms of social media platforms to manage and amplify his political campaign, thereby pursuing the ultimate objective of manipulating and influencing voters and shaping their electoral attitudes.

The Problem of “Non-Existing Deepfake” as a Regulatory Blind Spot

A significant issue that Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 – Artificial Intelligence Act (Regulation (EU) 2024/1689) deepens rather than resolves is the so-called “non-existing deepfake”, in which the generated synthetic content does not constitute the manipulation of an already existing person, voice, or event, but instead creates entirely new, non-existent subjects, objects, and factual situations. In such cases, there is no primary reference to reality, i.e., there is no authentic content that is being manipulated and against which a verification of authenticity can be carried out. This places this type of synthetic content within a so-called “dark zone” of regulation, as it remains outside the effective scope of the regulatory mechanism.

According to Article 3 (60) of the Artificial Intelligence Act, a “deepfake” means generated or manipulated image, audio, or video content that resembles existing persons, objects, places, entities, or events and falsely appears to a person to be authentic or truthful. This legal definition demonstrates that the normative scope is intrinsically linked to the existence of a real-world referent, thereby further excluding from its application those scenarios involving entirely synthetic (Łabuz 2025) , “non-existing” content.

Pursuant to Article 50 (2) of the AI Act, providers of systems generating synthetic audio, video, or images are required to ensure that such outputs are “marked in a machine-readable format and are detectable as artificially generated or manipulated.” In parallel, Article 50 (4) imposes a direct obligation on users of deepfake systems to disclose that the content has been artificially generated. The problem, however, remains that synthetic content lacking any authentic referent falls outside the effective reach of these regulatory mechanisms (Krastev 2024, pp. 285–292).

An interesting example of the relevance of transparency obligations under the Artificial Intelligence Act is the case of the AI-generated automated call impersonating the voice of Joe Biden during the 2024 New Hampshire primary election (FCC 2024). Although the case concerns the electoral process in the United States and falls outside the territorial scope of EU law, it clearly illustrates the type of risk that Article 3(60) and Article 50 of the AI Act seek to address: synthetic audio content that resembles the voice of an existing person and creates a false impression among recipients that it is authentic. In that case, an artificial intelligence system was used to generate the voice of then-President Joe Biden in a voice message urging voters not to vote. In addition, an algorithmic and spoofing mechanism was used to amplify the effect and reach the maximum possible number of participants in the electoral process. The problem is therefore not limited to the existence of disinformation as such, but also concerns the lack of clear disclosure that the content was synthetic and the inability of voters to recognize in a timely manner that they were exposed to AI-generated content that was not authentic. This is precisely why the AI Act shifts the focus from subsequent sanctioning to prior transparency, machine-readable marking, and explicit disclosure of the synthetic nature of content where such content may create a false impression of authenticity.

Conclusion

The legislative approach, both in the United States and at the supranational level of the European Union, demonstrates limited effectiveness in countering this phenomenon. The American model, based on § 20010 of the California Elections Code, is functionally constrained by the high standards of proof and by the immunity under Section 230 of the CDA, which in practice excludes the liability of platforms. The European approach, embodied in the Digital Services Act and the Artificial Intelligence Act, provides a broader and more proactive regulatory framework, yet it likewise fails to adequately address the challenges related to evidentiary standards and the technological tools required for the detection and identification of synthetic content.

A particular shortcoming of the Artificial Intelligence Act is that it leaves in a “dark zone” the so-called “non-existing deepfake”, where no real referent exists against which

authenticity can be verified, thereby limiting the possibility of effective control despite the high potential of such content to influence public attitudes. The absence of reliable, standardized, and widely applicable technological means for identifying deepfake content undermines the effectiveness of legal mechanisms, regardless of the regulatory approach adopted.

In this context, effective counteraction of deepfakes in the electoral process requires a reconsideration of the legislative approach and the development of an integrated model combining legal, technological, and institutional measures, as well as the introduction of clear evidentiary standards (Kirkov 2022, pp. 5–9) and the possibility of reliable technological verification of synthetic content. The main challenge is not limited to the lack of sufficient codification concerning the legal regulation of synthetic content, but lies in the structural lag of the regulatory framework behind technological development. This requires the establishment of a more flexible, preventive, and technologically neutral legal model for countering synthetic content – deepfakes, that directly affects the electoral process.

References

- RAJPUT, M., 2026. The Critical Threat of Deepfakes: Vulnerability and Resilience in Democratic Elections. *International Journal for Multidisciplinary Research*, vol. 8, no. 1. Available from: <https://doi.org/10.36948/ijfmr.2026.v08i01.63435>. [Viewed 2026-03-02].
- DAN, V., 2025. Deepfakes as a Democratic Threat: Experimental Evidence Shows Noxious Effects That Are Reducible Through Journalistic Fact Checks. *The International Journal of Press/Politics*. Available from: <https://doi.org/10.1177/19401612251317766>. [Viewed 2026-03-04].
- MICHALKIEWICZ-KĄDZIELA, E., 2024. The Impact of Deepfakes on Elections and Methods of Combating Disinformation in the Virtual World. *Teka Komisji Prawniczej PAN Oddział w Lublinie*. Available from: <https://doi.org/10.32084/tkp.8615>. [Viewed 2026-03-06].
- CALIFORNIA LEGISLATURE, Elections Code – ELEC § 20010. Online. Available from: <https://codes.findlaw.com/ca/elections-code/elec-sect-20010>. [Viewed 2026-03-08].
- CITRON, D. K. and WITTES, B., 2017. The Internet Will Not Break: Denying Bad Samaritans Section 230 Immunity. *Fordham Law Review*, vol. 86, no. 2, pp. 401–456. Available from: https://scholarship.law.bu.edu/faculty_scholarship/613. [Viewed 2026-03-10].
- NIKOLOV, B. and NIKOLOV, D., 2024. A scenario-based approach to cybersecurity training for seafarers. In: *Proceedings of the International Association of Maritime Universities Conference 2024*. Massachusetts Maritime Academy, United States, pp. 385–390.
- EUROPEAN PARLIAMENT and 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 – DSA). Available from: <https://eur-lex.europa.eu/eli/reg/2022/2065/oj/eng>. [Viewed 2026-03-12].
- FATHAIGH, R.; BUIJS, D. and VAN HOBOKEN, J., 2025. The Regulation of Disinformation under the Digital Services Act. *Media and Communication*. Available from:

<https://doi.org/10.17645/mac.9615>. Digital Services Act available from: <https://eur-lex.europa.eu/eli/reg/2022/2065/oj/eng>. [Viewed 2026-03-14].

VARBANOV, G., 2025. Application of deep learning in artificial intelligence systems for cyberattack identification and prevention. In: Proceedings of the 16th International Scientific and Practical Conference "Environment. Technology. Resources". Rezekne, Latvia, vol. V, pp. 323–327.

CONSTITUTIONAL COURT OF ROMANIA, 2024. Hotărârea nr. 32/2024 privind anularea procesului electoral cu privire la alegerea Președintelui României din anul 2024. Online. Available from: <https://lege5.ro/gratuit/ge3danjsm4ti/hotararea-nr-32-2024-privind-anularea-procesului-electoral-cu-privire-la-alegerea-presedintelui-romaniei-din-anul-2024>. [Viewed 2026-03-16].

EUROPEAN PARLIAMENT and 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). Available from: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>. [Viewed 2026-03-19].

LABUZ, M., 2025. A Teleological Interpretation of the Definition of DeepFakes in the EU Artificial Intelligence Act – A Purpose-Based Approach to Potential Problems With the Word "Existing". Policy & Internet, vol. 17. Available from: <https://doi.org/10.1002/poi3.435>. [Viewed 2026-03-21].

KRASTEV, D., 2024. Deepfake technology as a threat to information security in business. In: Law and Business in Contemporary Society. Varna, Bulgaria: University of Economics–Varna, pp. 285–292 [Viewed 2026-03-21].

FEDERAL COMMUNICATIONS COMMISSION, 2024. FCC fines man behind election interference scheme \$6 million for sending illegal robocalls that used deepfake generative AI technology. Online. 26 Sept. 2024. Available from: <https://www.fcc.gov/document/fcc-issues-6m-fine-nh-robocalls>. [Viewed 2026-03-24].

KIRKOV, A., 2022. Sudebno-ekspertnaya deyatel'nost' ekspertov v Bolgarii. Organizatsiya i problemy. In: Tsifrovizatsiya deyatel'nosti sudov: tekushchie i perspektivnye zadachi, pp. 5–9. [Viewed 2026-03-27].

THE USE OF DEEPPAKE TECHNOLOGIES FOR INFLUENCING ELECTORAL PROCESSES AND THEIR IMPACT ON NATIONAL SECURITY

Abstract. The study examines the technological phenomenon of deepfake as a tool for influencing electoral processes and its impact on national security through the formation and manipulation of attitudes of individuals and social groups toward the electoral process. In the context of increasing digitalization of society and the growing influence of social media, synthetically generated content is gaining significant potential to shape public opinion, particularly during sensitive periods such as election campaigns. The article analyzes the technological characteristics of deepfake technologies and their primary applications in undermining trust in electoral processes and in the actors involved, as well as the specific mechanisms through which voter attitudes are influenced. Particular attention is given to the risks to national security, including the erosion of public trust, the potential for external interference through information operations, and the capacity for destabilization of the social environment. The study also examines the challenges associated with the legal regulation of deepfake technologies, with reference to relevant European initiatives and the existing national legal framework. It concludes by arguing for the necessity of a comprehensive approach incorporating legal, technological, and institutional measures aimed at limiting the misuse of synthetic content and ensuring the security and integrity of electoral processes.

Keywords: history of the Bulgarian transport system; Bulgarian governments; railway transport

Assoc. Prof. Dr Gergana Varbanova

ORCID iD: 0000-0001-8122-4353

Web of Science ResearcherID: HKP-1334-2023

Nikola Vaptsarov Naval Academy

9000 Varna, Bulgaria

E-mail: g.varbanova@naval-acad.bg