

Criminalising Electoral Evidence: A Comparative Legal Framework for Electoral Integrity and Political Speech

Muhammad Fadel Kadir¹

¹ Fakultas Hukum, Universitas Muslim Indonesia, Indonesia

Correspondence: fadelkadir13@gmail.com

Article Info

Article history:

Received July 10th, 2026

Revised July 25th, 2026

Accepted July 29th, 2026

Keyword:

Criminalisation; Election deepfakes;
Electoral integrity; Political speech;
Synthetic evidence.

ABSTRACT

Existing regulation of election deepfakes had focused on the technology used or the falsity of content, leaving the boundary between electoral deception and political speech insufficiently defined. This study aimed to formulate a narrow basis for criminalising synthetic media used as false evidence in elections without expanding state authority to determine political truth. A normative legal method was applied through statutory, conceptual, and comparative analysis of legislation and judicial decisions from Indonesia, the European Union, South Korea, California, and Washington. The findings showed that deepfake was an unsuitable object of criminalisation because it described a production technique rather than the legal wrong involved. A more precise object was synthetic evidence, namely media created or materially altered to appear as an authentic record of an electoral fact that did not occur or differed materially from reality. The comparison showed that disclosure, temporal limits, fault requirements, and non-criminal remedies performed distinct functions. Criminal liability was justified only when an actor knew that a recording was false, intended it to be accepted as authentic to influence electoral choice or disrupt election administration, falsified a material fact, and created a concrete danger through its distribution. Clearly disclosed synthetic media remained within protected political speech.



© 2026 The Authors. Published by Punggawa Legacy Center. This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>)

INTRODUCTION

Synthetic media has changed the evidentiary appearance of political deception. Generative systems can now place a candidate's face or voice inside a recording that appears to document an event, yet *deepfake* remains a broad technological label rather than a stable legal category. It covers malicious impersonation, disclosed satire, fictional reconstruction, campaign experimentation, and other uses that do not carry the same wrong. Reviews of the field support that distinction (Kietzmann et al., 2020), while recent empirical mapping shows that *deepfake* research remains diverse across applications and disciplines (Wang, 2026). Election *deepfakes* can affect political attitudes under particular conditions (Dobber et al., 2021), but newer evidence indicates that they are not necessarily more credible than other false media and may instead weaken trust in authentic material (Barari et al., 2025). Those risks do not justify criminalising every synthetic political message. Simple editing can also create convincing false evidence, while openly fictional synthetic media may remain protected expression. The relevant legal question is therefore not whether a recording is technologically artificial, but whether it has been made to function as evidence of an electoral fact.

Political communication is an especially difficult setting for criminalisation. Campaigns contain exaggeration, selective presentation, emotional persuasion, parody, and claims whose truth cannot always be settled by a court without entering the substance of political contestation. Overbroad offences against false or misleading information can give public authorities a general power to police political truth. The danger is not confined to conviction. Investigation, content removal, and the threat of prosecution may suppress speech before a court decides whether the recording was false or deceptive. Election-focused scholarship has accordingly stressed the need to address *deepfake* harms without disregarding the ethical interests of voters, candidates, and electoral institutions (Diakopoulos &

Johnson, 2021). Freedom-of-expression analysis likewise shows that harmful *deepfakes* do not automatically fall outside the ordinary protection given to communication merely because they are synthetic (Barber, 2023). *Deepfake* regulation must consequently distinguish between the falsity of a proposition and the fabrication of documentary support for that proposition. The former may remain part of ordinary political dispute, while the latter adds a different wrong when a fabricated record is deliberately offered as authentic proof.

Existing rules reveal no single regulatory answer. The European Union's Artificial Intelligence Act relies primarily on disclosure for *deepfakes*, preserving circulation while requiring the artificial character of the content to be revealed. South Korea uses a ninety-day prohibition before polling day and allows labelled synthetic campaign media outside that period. California combines temporal limits, heightened fault, disclosure, and civil remedies, while Washington gives candidates injunctive and damages actions against synthetic electioneering communications. Indonesia has not enacted a specific election-*deepfake* offence. Comparative criminal-law research similarly identifies Indonesia's reliance on broader cyber and penal provisions rather than a dedicated electoral *deepfake* rule (Rahman & Anggriawan, 2025). The Constitutional Court has nevertheless required original and recent candidate photographs that are not excessively manipulated by artificial intelligence, linking authentic representation to voters' access to accurate information and to free, honest, and fair elections. Other Constitutional Court decisions have invalidated or narrowed vague information offences. Taken together, these materials expose a central tension: electoral authenticity may require legal protection, but that protection cannot be built from elastic concepts such as misleading content, excessive manipulation, or online disorder without more precise limits.

The literature has examined the technical production of *deepfakes*, their epistemic effects, and the ethical responsibilities of platforms, campaigns, and audiences. Broader research on intentionally deceptive information has also shown that categories of falsity differ in their production, presentation, and claimed relationship to evidence (Damstra et al., 2021). Legal scholarship has separately examined the criminalisation and control of manipulated images, showing why legal intervention cannot be reduced to technological detection alone (Kirchengast, 2020). A narrower question remains insufficiently developed: what exactly should criminal law treat as the wrongful object when synthetic media is used in an election, and which cumulative conditions are required before that use becomes punishable? Technology-based definitions obscure the function of the content. General falsehood offences risk criminalising ordinary political speech. Disclosure-only models may also be inadequate when concealment of the artificial source is the means by which the deception operates. The research gap lies in the absence of a doctrinal model that separates synthetic media as a form of expression from *synthetic evidence* as fabricated documentary support and then calibrates liability to fault, electoral materiality, and concrete danger.

This study aims to formulate that model. It argues that criminal law should apply only when artificially created or materially altered media is knowingly presented as an authentic record of a material electoral fact, with the specific purpose of influencing electoral choice or disrupting election administration, and in circumstances that create a concrete danger. The concept of *synthetic evidence* is introduced to identify this object without tying liability to a particular technology. Comparative rules are then examined not as complete national models, but as combinations of authenticity requirements, disclosure duties, temporal restrictions, fault standards, and non-criminal remedies. The analysis leads to a tiered response. Clearly disclosed synthetic political expression remains lawful. Defective disclosure that lacks serious deception should be handled through correction or other non-criminal measures. Criminal punishment is reserved for deliberate fabrication that uses the authority of an apparently authentic record to endanger the reliability of documentary information on which voters and election administrators depend.

RESEARCH METHODS

This study is a normative legal inquiry into the limits of criminal law in regulating synthetic media during elections. It uses statutory, conceptual, and comparative approaches. Normative legal research is directed to the identification, interpretation, and evaluation of legal norms rather than measurement through field research, an orientation also recognised in contemporary Indonesian legal-method scholarship (Negara, 2023). The statutory approach examines legislation and judicial decisions concerning campaign authenticity, disclosure duties, restrictions on distribution, and legal

responsibility. The conceptual approach distinguishes synthetic media that remains political speech from *synthetic evidence* presented as an authentic record. The comparative approach evaluates regulatory variations in Indonesia, the European Union, South Korea, California, and Washington without examining each legal system as a whole. Primary materials consist of legislation and judicial decisions; secondary materials include journal articles, books, institutional reports, and policy documents. All materials were collected through library research using official institutional websites, legal databases, and academic publications available in full text.

The legal materials are analysed qualitatively through legal interpretation and comparison. Each rule and judicial consideration is read in relation to the legal interest protected, the conduct prohibited, the required form of fault, and the legal consequence imposed. The comparison proceeds without treating foreign rules as interchangeable or detached from their legal setting, consistent with recent scholarship that calls attention to what legal comparison itself assumes about similarity and difference (Petretta, 2020). The results are compared to assess the clarity of the prohibition, the relationship between fault and electoral harm, and the proportionality of the response. Elements that are too broad or that invite authorities to judge political truth are separated from those needed to protect electoral authenticity. This integration of doctrinal interpretation with conceptual evaluation is also consistent with work that treats doctrinal scholarship as a distinct but complementary form of legal inquiry (Taekema, 2021).

RESULTS AND DISCUSSION

1. From Synthetic Media to Synthetic Evidence

A satirical video and a false announcement impersonating an election authority can be produced with the same software, but their technical similarity does not give them the same legal character. The former may be understood as criticism or fiction, while the latter asks the public to believe that an institution actually issued the message. Technical surveys confirm that *deepfakes* comprise multiple forms of generated and manipulated media whose production characteristics do not determine their social or legal function (Mirsky & Lee, 2021). The term *deepfake* is therefore inadequate as the object of criminalisation. This article instead uses *synthetic evidence* for media created or materially altered so that it appears to document an electoral fact that did not occur or occurred in a materially different way. Evidence is used here in a public, not solely procedural, sense: it refers to a recording offered as a reason to believe a fact even when it is never introduced in court. Intent is not built into the definition. Knowledge and deceptive purpose determine responsibility later; they do not determine the character of the content.

Technology-dependent prohibitions produce two errors at once. A rule may fail to reach a simple edit that changes the meaning of a statement because the file was not created with the named artificial-intelligence technique. Research comparing cheap and technologically sophisticated manipulation also indicates that political deception cannot be understood solely through the complexity of the production method (Hameleers, 2024). The same rule may capture a highly realistic fictional work even though its audience understands that the scene never occurred. The EU definition, which asks whether manipulated media falsely appears authentic, moves attention from the tool to the way content is perceived (European Parliament & Council, 2024). Yet technological neutrality is not sufficient when terms such as manipulated media or misleading content remain unbounded. Such language may still extend to criticism, ordinary error, or creative expression. A workable rule must ask whether the content appears to record an actual event, whether the represented fact is electorally material, and whether the actor knew of the falsity. These questions should not be collapsed into a technical definition. Their separation allows the law to remain applicable as production techniques change while preventing technical sophistication from becoming the measure of criminality.

The distinction becomes clearer when an allegation is converted into a recording that appears to support it. A statement that a candidate accepted a bribe still invites readers to ask for its source. A synthetic video depicting the transaction gives a different impression: a camera appears to have witnessed the event. The appearance of authenticity does not always depend on flawless visual quality. The identity of the account, the caption, or the manner of distribution may make a relatively simple recording appear official. Research on political *deepfakes* similarly shows that doubt and perceived plausibility are shaped by the content and the surrounding communicative context, not only by technical realism (Hameleers et al., 2024). An authenticity claim can therefore be explicit or arise implicitly from

context. California's reasonable-viewer formulation is useful, although comparison with an original file will not always be possible because some recordings are created entirely from scratch. The central inquiry is whether the media is offered as documentation of an event that is nonexistent or materially distorted. The actor has then done more than state a falsehood. The actor has supplied false documentary support so that the falsehood appears already proved.

Falsity concerns the content of a claim. Deception concerns the way an actor attempts to produce a mistaken belief, and false evidence is one means of doing so. These concepts do not always form a fixed hierarchy. Research on misinformation also cautions that falsity must be assessed against available expertise and evidence, while deceptive intent remains a separate question (Vraga & Bode, 2020). A person may lie without fabricating a recording, while a false recording may be redistributed by someone unaware of its origin. The distinction matters because criminal law should not be directed at every incorrect political statement. Harsh evaluation, failed prediction, and factual error require a different response from the deliberate use of a fabricated record as documentation of an actual event. The object of regulation must therefore remain separate from the actor's fault and from the degree of danger later created. Criminal punishment becomes relevant only where those additional elements place the conduct beyond ordinary political dispute.

The harm does not necessarily wait until a fabrication is fully believed. Vaccari and Chadwick (2020) found that deceptive political *deepfakes* could increase uncertainty even when belief in the false statement did not significantly rise. Their experiment does not establish that *deepfakes* decide votes or electoral outcomes. It shows a narrower point: uncertainty can itself matter. A related problem is that authentic evidence may become easier to deny once synthetic media is widely plausible. The normative risk may therefore arise either from misplaced belief or from the erosion of confidence in genuine documentation.

The state does not need authority to determine every political truth in order to address this problem. A narrower legal interest is the minimum reliability of documentary information used to form electoral choices and to administer elections. The epistemic threat of *deepfakes* lies partly in their capacity to weaken the evidential role of recordings, including genuine ones (Fallis, 2021). Voters need a reasonable opportunity to distinguish authentic records from fabrications. Election authorities must also be protected from false recordings that appropriate their identity or voice. This formulation does not make candidate reputation the sole interest (Kadir, 2026b, 2026a). A synthetic announcement impersonating an election institution may be more damaging than a video that merely embarrasses a candidate. At the same time, offensive or unfair content does not become *synthetic evidence* unless it is presented as documentation of an actual fact. Protection remains tied to documentary information with a meaningful electoral connection.

Clear disclosure can weaken the appearance that synthetic media is an actual record. Recognisable satire and simulations explained from the outset also remain within political speech. Experimental research on political *deepfake* parodies shows that AI disclaimers can improve recognition of parody, although their effect depends on audience comprehension (Lu & Yuan, 2024). Disclosure, however, cannot be treated as conclusive in every setting. A mark may be obscured, contradicted by the accompanying narrative, or become unclear after redistribution. Its significance must be assessed together with the overall presentation. Synthetic media is not punishable merely because it is realistic or election-related. The legal problem becomes sharper when it is offered as authentic documentation of a material fact. Even then, the conceptual classification only identifies the possible object of regulation. Criminal liability still depends on the actor's fault, the danger created, and whether less restrictive remedies are available.

2. Comparing Regulatory Techniques and Their Constitutional Trade-offs

Decision No. 166/PUU-XXI/2023 of the Constitutional Court of Indonesia did not create a general rule for synthetic campaign media. The Court conditionally interpreted the statutory phrase referring to a candidate's self-image only insofar as it concerned photographs or images. Those materials must be original, recent, and not excessively manipulated through artificial intelligence. The application had requested a wider scope, including voice and combined audio-visual material, but the operative order did not grant that part. The reasoning remains important. The Court connected excessive manipulation with voters' ability to make informed decisions and with free, honest, and fair elections. That reasoning can support a broader doctrinal concern with electoral authenticity, but the extension

belongs to this article rather than to the binding order itself. The phrase excessive manipulation is also unsafe as a criminal element. Without a test of material alteration and deceptive presentation, it gives enforcement authorities too much room to decide which edits are acceptable. The judgment supplies a constitutional interest, not a ready-made deepfake offence (Constitutional Court of Indonesia, 2025b).

A different boundary emerges from Decision No. 78/PUU-XXI/2023 and Decision No. 115/PUU-XXII/2024. The first invalidated Articles 14 and 15 of Law No. 1 of 1946, while the second confined the statutory term public disorder to disturbance in physical space rather than digital commotion. Neither decision means that every information offence must require completed physical damage. Their more modest lesson concerns precision. When falsity or social disturbance supports criminal liability, the relevant conduct and consequence must be reasonably foreseeable. A synthetic-evidence offence cannot rest only on words such as false, misleading, or excessive. Political criticism that causes online controversy should not be equated with a deceptive recording that places election administration or voter choice in concrete danger (Constitutional Court of Indonesia, 2024, 2025a).

Article 50 of the EU Artificial Intelligence Act chooses disclosure rather than a general prohibition. A deployer of an AI system that creates or manipulates a *deepfake* must disclose the artificial nature of the content. Clearly artistic, fictional, or satirical works retain room for a form of disclosure that does not unduly interfere with their presentation. This is a cross-sector AI rule, not an election offence. Analysis of the Act's *deepfake* provisions shows that transparency obligations are central to its treatment of synthetic media (Łabuz, 2024). Its relevance here lies in the fact that the disclosure duty does not depend on a state determination that the political opinion expressed is false. The audience receives information about how the media was produced, while the political message remains open to debate. The intervention concerns provenance rather than the truth of the political position (European Parliament & Council, 2024).

Disclosure is most useful when the central problem is that the audience does not know that the content is artificial. It is weaker when impersonation of an official source is the means of deception. A person mimicking an election authority is not merely omitting a label but constructing a false source. Disclosure regimes also depend on the mark remaining visible and intelligible as content is redistributed. This is a structural limit, not an empirical claim that the EU system has failed. Disclosure is nevertheless proportionate as an initial rule because it does not ban synthetic expression. Once concealment is knowing and directed toward electoral harm, however, analysis must shift to fault and danger. In that setting, absence of disclosure is part of the deceptive design rather than a minor technical defect. Human-rights analysis similarly stresses that controls on harmful *deepfakes* require safeguards against unnecessary restrictions on expression (Romero Moreno, 2024).

South Korean law gives time a different role. Article 82-8 of the Public Official Election Act prohibits the production and circulation of AI-generated campaign media that is difficult to distinguish from reality during the ninety days before polling day. Outside that period, comparable content may be used if its artificial character is clearly marked. Recent discussion of political *deepfake* advertising also uses South Korea to illustrate how synthetic campaigning can combine legitimate communicative uses with risks to the quality of democratic deliberation (Łabuz & Strnad, 2025). The temporal line is precise and recognises that opportunities for verification and correction narrow as voting approaches. Yet timing does not establish deception. A labelled simulation and a false recording posted by an account resembling an election institution remain different even if released on the same day. Proximity to polling is therefore useful as a factor increasing danger, but a blanket temporal ban is too coarse to serve as the sole basis of criminalisation (Kadir, 2026c).

California and Washington should not be treated as one model concerned only with candidate reputation. California Elections Code section 20010 applies, within sixty days of an election, to materially deceptive audio-visual media of a candidate distributed with *actual malice* and an intent to injure reputation or deceive voters. Disclosure, satire, and parody limit its application. Section 20012, introduced later, extends beyond candidate reputation to fabricated depictions of election officials, elected officials, voting equipment, and other election-related matters that are reasonably likely to undermine confidence in an electoral result. It permits injunctive and damages actions by several affected parties. Washington remains more candidate-centred, granting candidates equitable relief and damages where their appearance, speech, or conduct is altered in an electioneering communication; a compliant disclosure is an affirmative defence. The American concept of *actual malice* cannot simply be imported into Indonesian criminal doctrine. Its functional lesson is narrower: liability should require

knowledge of falsity or an equivalently demanding fault threshold, together with an electoral purpose. These statutes also demonstrate that rapid civil relief and damages can be used without making imprisonment the first response (California Legislative Information, n.d.; Washington State Legislature, 2023).

The shared weakness is not lenient punishment. Authenticity, disclosure, temporal limits, fault, and civil remedies often overlap rather than appearing in separate regulatory boxes, and scholarship likewise treats *deepfakes* as a problem that cuts across existing legal categories (van der Sloot & Wagenveld, 2022). No single technique can manage every form of manipulation. Disclosure cannot replace fault, a temporal limit does not prove deceptive purpose, and a candidate-centred right of action may leave election administration insufficiently protected. Conversely, a prohibition covering every altered recording would erase the difference between creative criticism and fabricated evidence. Increasing penalties cannot repair an unclear object or a missing fault element. More generally, the EU debate shows that AI regulation must negotiate regulatory tools and fundamental-rights constraints rather than assume that one logic can perform both tasks (Almada & Petit, 2025). A design error may instead widen the possibility of selective enforcement against political opponents during a contested campaign.

Indonesian law can use the limited function of each technique without copying its foreign form. Constitutional Court reasoning supplies an interest in honest electoral representation and a warning against flexible information offences. Disclosure is an appropriate initial rule because synthetic media that openly reveals its nature should not be prohibited. Recent work on AI-generated political visualisations likewise shows that synthetic campaigning includes both legitimate communicative uses and manipulative practices, which makes a blanket ban difficult to justify (Pawelec & Łabuz, 2026). Proximity to polling may strengthen the assessment of danger without creating liability by itself. Correction, civil relief, or other non-criminal responses remain necessary because not every disclosure failure warrants punishment. Each technique must perform a limited task; none should be forced to protect the entire electoral process. Such an arrangement is more defensible than a broad prohibition that leaves most decisions to criminal enforcement authorities.

The unresolved question is the point at which criminal law becomes justified. That point is not reached merely by the use of artificial intelligence or by disagreement over the truth of a campaign message. It approaches only when synthetic media is used as an authentic record, the actor knows that it is false, and its distribution creates a concrete danger to voter choice or election administration. Disclosure, time, and non-criminal remedies operate around that core. They can weaken an authenticity claim, intensify risk, or provide a less restrictive response. Current research also cautions against treating the persuasive power of generative political visuals as self-evident (Kruschinski & Votta, 2026). A criminal rule must therefore emerge from a narrow combination of elements rather than from general anxiety about a technology whose uses are not uniformly harmful.

3. Criminalising Synthetic Electoral Evidence Without Policing Political Speech

A general prohibition on synthetic media cannot be justified merely because the technology can create realistic recordings. The legal problem begins when an artificial record is used to damage the reliability of documentary information on which voters and election administrators depend. That is the interest requiring protection, not a state power to determine all political truth. The Constitutional Court's reasoning on excessive manipulation of candidate images can be developed in this direction, although the judgment itself did not establish a general principle for all synthetic media. Its other information-law decisions also show why a criminal rule cannot depend on elastic terminology. Recent analysis of Indonesia's new Penal Code similarly argues that false-news offences require specific and strict limits if criminalisation is to remain compatible with freedom of expression (Prahassacitta & Harkrisnowo, 2024). The less precise the object, the easier it becomes for a campaign dispute to be converted into a criminal complaint. Criminal punishment should therefore be considered only after the prohibited conduct is clearly distinguishable from satire, ordinary error, or harsh criticism. This subsidiary role also reflects the broader criminalisation question of which wrongs are sufficiently apt for criminal-law condemnation rather than merely undesirable conduct (Tadros, 2020). Correction and non-criminal remedies remain prior options unless the harm reaches the threshold required for the most severe state response.

Consider a video that appears to record a candidate accepting money even though the entire scene was produced digitally (Kadir, 2025). It is *synthetic evidence* when the media has been created or materially altered so that it appears to document an electoral fact that did not occur or occurred with a materially different meaning. This definition concerns the character of the content, not the fault of every person who encounters it. The same recording remains *synthetic evidence* when forwarded by someone who believes it is genuine. Creators and distributors should therefore be subject to the same standard: responsibility arises only when they know that the recording is false or materially altered and use it so that it will be accepted as authentic documentation. Production alone is not enough, and awareness that a file was generated with AI is not equivalent to knowledge that it falsely represents an event. Other participants who commission or control dissemination remain subject to ordinary doctrines of participation rather than a new catalogue of digital actors.

The same recording may be held by a journalist, a propagandist, or a user unaware of its origin. Knowledge alone is therefore insufficient. The actor must also intend the public to accept the recording as authentic in order to influence electoral choice or disrupt election administration. Both elements are cumulative. A special purpose is necessary because identical content may be used to expose a deception rather than to carry it out. California law illustrates the protective function of a demanding fault threshold through *actual malice*, but Indonesian law does not need to adopt that term (Kadir, 2024). A more appropriate construction is intent that includes knowledge of falsity and a specific purpose to deceive. Use of an account resembling an official institution or removal of an existing disclosure may support an inference, but neither proves intent automatically. The conclusion must arise from the circumstances as a whole. This threshold prevents ordinary negligence from being equated with a planned electoral deception.

Merely naming a candidate does not make a manipulation electorally material. Materiality depends on the importance of the falsified fact. The recording must concern information reasonably capable of affecting voter choice or the reliability of election administration. A fabricated position on a central policy question may meet that threshold, as may synthetic audio impersonating an election official and giving false voting instructions. Minor alteration that does not change political or procedural meaning should remain outside criminal law. Prosecutors need not prove that the final election result changed, an evidentiary burden that would usually be impossible. The electoral connection cannot, however, be assumed simply because the content appears during a campaign. Materiality concerns the significance of the falsified information, not the fame of the person depicted or the volume of public attention it receives.

Materiality does not establish concrete danger. A simulation may concern an important fact without creating a risk because its artificial character is clear from the outset. Danger arises from a distribution context in which the recording is genuinely capable of disrupting voter choice or election administration. Proximity to polling may increase that risk because time for checking and correction narrows. Timing nevertheless does not replace knowledge, deceptive purpose, or materiality. The proposed model therefore requires concrete danger so that criminal liability is not based on an assumption that all synthetic media is harmful. The danger must be demonstrated through the facts of distribution, not inferred solely from the technical capacity of the medium.

Disclosure performs a different function from punishment. In this model, a label is evidence that content was not intended as an actual record rather than an automatic immunity. A clear statement can weaken the appearance of authenticity, while a hidden statement or one contradicted by the accompanying narrative may not be sufficient. Experimental research on warnings and fact-check tags shows that the wording and form of corrective signals affect how users process questionable content (Clayton et al., 2020). The legal question remains whether the artificial character of the content is clearly understandable in its distribution context. A disclosure failure without deceptive intent or serious danger should ordinarily be addressed through correction, reply, or another non-criminal remedy. Research on misinformation shows both the value of correction and the difficulty of fully reversing a false belief after exposure (Ecker et al., 2022). This functional approach avoids profession-based exemptions: research, reporting, satire, and ordinary citizen use remain outside the offence when the content is not used to make the public believe that it documents an actual event.

A finding that a recording was manipulated should not depend on a single detection tool. Human observers often overestimate their ability to identify *deepfakes* (Köbis et al., 2021). Comparative experiments also show that neither human judgment nor automated detection is uniformly reliable

across cases (Groh et al., 2022). Media-forensics research therefore treats detection as a multi-method problem rather than a single-score decision (Verdoliva, 2020). Forensic assessment must remain open to testing and be considered with other evidence concerning the origin and use of the content. The model's normative boundary can then be stated concisely. Criminal liability may be considered only when all elements are present: the actor knows that the recording is false or materially altered, intends it to be accepted as authentic in order to influence electoral choice or disrupt election administration, falsifies a material electoral fact, and creates a concrete danger through the context of distribution. This is not a final statutory clause. It is a minimum threshold designed to protect documentary reliability without turning criminal law into a general prohibition on political speech. Legislative implementation would still need to determine competent institutions, procedure, and proportionate sanctions.

CONCLUSION

The concept of *deepfake* is an unsuitable foundation for criminalisation because it identifies a media form or production technique. The legal problem arises when created or altered media is used as an authentic record of an electoral fact. This article describes that object as *synthetic evidence*. The comparison shows that authenticity, disclosure, timing, fault, and non-criminal remedies cannot operate alone. Disclosure preserves space for satire, criticism, and creative use, but it is inadequate when impersonation of an official source is itself the method of deception. Temporal proximity may increase risk by narrowing the opportunity for correction, yet closeness to polling does not establish an offence by itself.

Criminalisation may be considered only when every element is satisfied. The actor must know that the recording is false or materially altered and intend it to be accepted as authentic in order to influence voter choice or disrupt election administration. The falsified fact must be material, and the distribution context must create a concrete danger. Synthetic media that clearly reveals its artificial character remains within lawful political speech. Disclosure failures that lack deceptive intent or serious danger should instead be addressed through correction or non-criminal remedies. Democracy can survive synthetic evidence, but not through a general prohibition on artificial media. Criminal law must remain confined to fabricated recordings that genuinely deceive and endanger elections without granting the state a broad authority to determine political truth.

REFERENCES

- Almada, M., & Petit, N. (2025). The EU AI Act: Between the rock of product safety and the hard place of fundamental rights. *Common Market Law Review*, 62(1), 85–120. <https://doi.org/10.54648/COLA2025004>
- Barari, S., Lucas, C., & Munger, K. (2025). Political deepfakes are as credible as other fake media and (sometimes) real media. *The Journal of Politics*, 87(2), 510–526. <https://doi.org/10.1086/732990>
- Barber, A. (2023). Freedom of expression meets deepfakes. *Synthese*, 202, Article 40. <https://doi.org/10.1007/s11229-023-04266-4>
- California Legislative Information. (n.d.). California Elections Code §§ 20010–20012. Retrieved August 7, 2026, from https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?article=&chapter=1.&division=20.&lawCode=ELEC&part=&title=
- Clayton, K., Blair, S., Busam, J. A., Forstner, S., Glance, J., Green, G., Kawata, A., Kovvuri, A., Martin, J., Morgan, E., Sandhu, M., Sang, R., Scholz-Bright, R. A., Welch, A. T., Wolff, A. G., & Nyhan, B. (2020). Real solutions for fake news? Measuring the effectiveness of general warnings and fact-check tags in reducing belief in false stories on social media. *Political Behavior*, 42(4), 1073–1095. <https://doi.org/10.1007/s11109-019-09533-0>
- Constitutional Court of Indonesia. (2024). Decision No. 78/PUU-XXI/2023. https://mkri.id/public/content/persidangan/putusan/putusan_mkri_9709_1711005671.pdf
- Constitutional Court of Indonesia. (2025a). Decision No. 115/PUU-XXII/2024. <https://www.mkri.id/perkara/persidangan/putusan?search=115>
- Constitutional Court of Indonesia. (2025b). Decision No. 166/PUU-XXI/2023. https://mkri.id/public/content/persidangan/putusan/putusan_mkri_11343_1735801113.pdf
- Damstra, A., Boomgaarden, H. G., Broda, E., Lindgren, E., Strömbäck, J., Tsifti, Y., & Vliegenthart, R. (2021). What does fake look like? A review of the literature on intentional deception in the news and on social media. *Journalism Studies*, 22(14), 1947–1963. <https://doi.org/10.1080/1461670X.2021.1979423>

- Diakopoulos, N., & Johnson, D. (2021). Anticipating and addressing the ethical implications of deepfakes in the context of elections. *New Media & Society*, 23(7), 2072–2098. <https://doi.org/10.1177/1461444820925811>
- Dobber, T., Metoui, N., Trilling, D., Helberger, N., & de Vreese, C. H. (2021). Do (microtargeted) deepfakes have real effects on political attitudes? *The International Journal of Press/Politics*, 26(1), 69–91. <https://doi.org/10.1177/1940161220944364>
- Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N. M., Kendeou, P., Vraga, E. K., & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, 1, 13–29. <https://doi.org/10.1038/s44159-021-00006-y>
- European Parliament & Council. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence. *Official Journal of the European Union*. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>
- Fallis, D. (2021). The epistemic threat of deepfakes. *Philosophy & Technology*, 34, 623–643. <https://doi.org/10.1007/s13347-020-00419-2>
- Groh, M., Epstein, Z., Firestone, C., & Picard, R. W. (2022). Deepfake detection by human crowds, machines, and machine-informed crowds. *Proceedings of the National Academy of Sciences*, 119(1), e2110013119. <https://doi.org/10.1073/pnas.2110013119>
- Hameleers, M. (2024). Cheap versus deep manipulation: The effects of cheapfakes versus deepfakes in a political setting. *International Journal of Public Opinion Research*, 36(1), edae004. <https://doi.org/10.1093/ijpor/edae004>
- Hameleers, M., van der Meer, T. G. L. A., & Dobber, T. (2024). They would never say anything like this! Reasons to doubt political deepfakes. *European Journal of Communication*, 39(1), 56–70.
- Kadir, Z. K. (2024). Dari Dualisme ke Monisme: Transformasi Konsep Mens Rea dalam Kodifikasi KUHP di Negara-Negara Poskolonial. *Jurnal Litigasi Amsir*, (Special Issue), 142–155.
- Kadir, Z. K. (2025). Kejahatan Berbasis Identitas Digital: Menggagas Kebijakan Kriminal untuk Dunia Metaverse. *Jurnal Litigasi Amsir*, 12(2), 124–137.
- Kadir, Z. K. (2026a). Beyond Firearms: Mapping Global Patterns of Non-Firearm Homicide in Comparative Criminological Perspective. *Punggawa Global Research: Jurnal Multidisiplin*, 1(2), 1–10.
- Kadir, Z. K. (2026b). Preventing Murder Without Expanding Punishment: Rethinking Homicide Policy Beyond Deterrence. *Punggawa Law Review*, 1(3), 49–58. <https://doi.org/10.67707/plr.v1i3.40>
- Kadir, Z. K. (2026c). Rural Criminology and the Cultural Authorization of Homicide: How Lethal Violence Becomes Socially Permissible. *Punggawa Social Inquiry: Journal of Crime, Culture, and Social Dynamics*, 1(1), 1–10.
- Kietzmann, J., Lee, L. W., McCarthy, I. P., & Kietzmann, T. C. (2020). Deepfakes: Trick or treat? *Business Horizons*, 63(2), 135–146. <https://doi.org/10.1016/j.bushor.2019.11.006>
- Kirchengast, T. (2020). Deepfakes and image manipulation: Criminalisation and control. *Information & Communications Technology Law*, 29(3), 308–323. <https://doi.org/10.1080/13600834.2020.1794615>
- Köbis, N. C., Doležalová, B., & Soraperra, I. (2021). Fooled twice: People cannot detect deepfakes but think they can. *iScience*, 24(11), 103364. <https://doi.org/10.1016/j.isci.2021.103364>
- Kruschinski, S., & Votta, F. (2026). Efficiency tool or manipulation machine? Assessing dominant claims about the quantity, quality, and persuasive power of visual generative AI in political campaigns. *Zeitschrift für Politikwissenschaft*, 36, 79–89. <https://doi.org/10.1007/s41358-026-00457-w>
- Łabuz, M. (2024). Deep fakes and the Artificial Intelligence Act—An important signal or a missed opportunity? *Policy & Internet*, 16(4), 783–800. <https://doi.org/10.1002/poi3.406>
- Łabuz, M., & Strnad, G. (2025). Political advertising with the use of deep fakes—Can deliberative democracy really benefit from this strategy? *Politics & Policy*, 53(3), e70035. <https://doi.org/10.1111/polp.70035>
- Lu, H., & Yuan, S. (2024). “I know it’s a deepfake”: The role of AI disclaimers and comprehension in the processing of deepfake parodies. *Journal of Communication*, 74(5), 359–373. <https://doi.org/10.1093/joc/jqae022>
- Mirsky, Y., & Lee, W. (2021). The creation and detection of deepfakes: A survey. *ACM Computing Surveys*, 54(1), Article 7, 1–41. <https://doi.org/10.1145/3425780>
- National Law Information Center. (2026). Public Official Election Act, Article 82-8. <https://www.law.go.kr/LSW/lsLinkCommonInfo.do?ancYnChk=&chrClsCd=&lsJoLnkSeq=1022178609>
- Negara, T. A. S. (2023). Normative legal research in Indonesia: Its originis and approaches. *Audito Comparative Law Journal*, 4(1), 1–9. <https://doi.org/10.22219/aclj.v4i1.24855>

- Pawelec, M., & Łabuz, M. (2026). Are softfakes really “soft”? Semantic, ethical, and legal perspectives on deepfakes and AI-generated visualizations in political campaigning. *International Journal of Politics, Culture, and Society*. Advance online publication. <https://doi.org/10.1007/s10767-026-09579-z>
- Petretta, I. (2020). The question of comparison. *The American Journal of Comparative Law*, 68(4), 893–928. <https://doi.org/10.1093/ajcl/avab003>
- Prahassacitta, V., & Harkrisnowo, H. (2024). The criminalization of fake news: Critique on Indonesia’s new Penal Code. *Criminal Law Forum*, 35(1), 83–120. <https://doi.org/10.1007/s10609-024-09478-y>
- Rahman, R. A., & Anggriawan, R. (2025). Deepfake and election crimes: Comparative perspectives from Indonesia, India, Pakistan, and the U.S. *Indonesian Comparative Law Review*, 7(2), 132–146. <https://doi.org/10.18196/iclr.v7i2.26337>
- Romero Moreno, F. (2024). Generative AI and deepfakes: A human rights approach to tackling harmful content. *International Review of Law, Computers & Technology*, 38(3), 297–326. <https://doi.org/10.1080/13600869.2024.2324540>
- Tadros, V. (2020). Criminalization: In and out. *Criminal Law and Philosophy*, 14, 365–380. <https://doi.org/10.1007/s11572-020-09536-7>
- Taekema, S. (2021). Methodologies of rule of law research: Why legal philosophy needs empirical and doctrinal scholarship. *Law and Philosophy*, 40, 33–66. <https://doi.org/10.1007/s10982-020-09388-1>
- Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, 6(1), 1–13. <https://doi.org/10.1177/2056305120903408>
- van der Sloot, B., & Wagenveld, Y. (2022). Deepfakes: Regulatory challenges for the synthetic society. *Computer Law & Security Review*, 46, Article 105716. <https://doi.org/10.1016/j.clsr.2022.105716>
- Verdoliva, L. (2020). Media forensics and deepfakes: An overview. *IEEE Journal of Selected Topics in Signal Processing*, 14(5), 910–932. <https://doi.org/10.1109/JSTSP.2020.3002101>
- Vraga, E. K., & Bode, L. (2020). Defining misinformation and understanding its bounded nature: Using expertise and evidence for describing misinformation. *Political Communication*, 37(1), 136–144. <https://doi.org/10.1080/10584609.2020.1716500>
- Wang, S. (2026). Mapping the landscape of deepfake research: A systematic review of empirical studies. *Big Data & Society*, 13(1), 1–13. <https://doi.org/10.1177/20539517261434331>
- Washington State Legislature. (2023). Revised Code of Washington § 29B.35.010. <https://app.leg.wa.gov/RCW/default.aspx?cite=29B.35.010>