

Prohibiting Political Deepfakes in Election Campaigns Based on the Principles of Honest and Fair Elections

Fahmi Nur Wicaksono

Universitas Jenderal Soedirman

Correspondence: nur.wicaksono@mhs.unsoed.ac.id

Received: 29 May 2026

Accepted: 28 August 2026

Published: 31 August 2026

Abstrak

Artikel ini mengkaji kebutuhan larangan *deepfake* politik manipulatif dalam kampanye pemilu di Indonesia serta merumuskan model pengaturan yang sejalan dengan prinsip pemilu jujur dan adil. Penelitian hukum normatif ini menggunakan pendekatan perundang-undangan, konseptual, dan kasus untuk menganalisis kekosongan regulasi yang ada. Hasil penelitian menunjukkan bahwa instrumen hukum saat ini, seperti UU Pemilu, UU ITE, dan Peraturan penyelenggara pemilu, masih bersifat parsial. Regulasi tersebut belum memiliki mekanisme operasional khusus untuk menindak manipulasi identitas berbasis kecerdasan buatan (AI) yang secara nyata mengancam integritas dan otonomi pemilih. Sebagai implikasi dan jawaban atas permasalahan tersebut, penelitian ini mengusulkan kerangka hukum yang lebih proporsional. Model yang ditawarkan mencakup larangan spesifik terhadap pemalsuan identitas politik, kewajiban pelabelan konten AI, mekanisme klarifikasi dan penghapusan konten secara cepat, serta pedoman koordinasi kelembagaan antara KPU, Bawaslu, Kominfo, dan platform digital. Implementasi model ini berdampak pada penguatan kepastian hukum dalam merespons disinformasi digital secara terukur, sehingga kemurnian kehendak pemilih tetap terlindungi tanpa memberangus kebebasan berekspresi seperti satir yang sah.

Kata kunci: *deepfake* politik; kampanye pemilu; pemilu jujur dan adil; kecerdasan buatan; integritas pemilu.

Abstract

This article examines the need to prohibit manipulative political deepfakes in Indonesian election campaigns and formulates a regulatory model aligned with the principles of honest and fair elections. This normative legal research uses statutory, conceptual, and case approaches to analyze existing regulatory gaps. The results indicate that current legal instruments, such as the Election Law, the Electronic Information and Transactions (ITE) Law, and electoral management body regulations, remain partial. These regulations lack specific operational mechanisms to address artificial intelligence (AI)-based identity manipulation that significantly threatens electoral integrity and voter autonomy. As an implication and answer to these issues, this study proposes a more proportionate legal framework. The proposed model includes a specific prohibition on political identity falsification, mandatory AI content labeling, rapid clarification and content removal mechanisms, and institutional coordination guidelines among the KPU, Bawaslu, Kominfo, and digital platforms. The implementation of this model strengthens legal certainty in responding to digital disinformation in a measured manner, ensuring that the purity of voter

intent remains protected without suppressing freedom of expression, such as legitimate satire.

Keywords: *political deepfake; election campaign; honest and fair elections; artificial intelligence; electoral integrity.*

INTRODUCTION

Electoral democracy does not end with voting and vote counting. The quality of an election is also determined by the information environment that shapes citizens' political choices before they enter the voting booth. Habermas characterizes the public sphere as an arena in which political opinion and will are formed through rational, open, and unmanipulated communication. This framework is consistent with Article 22E(1) of the 1945 Constitution of the Republic of Indonesia, which requires elections to be direct, general, free, confidential, honest, and fair.¹

Electoral-integrity studies regard campaigning as a decisive stage because voters assess candidates through the information they receive, share, and trust. Indonesian literature on hoaxes, buzzers, and political disinformation shows that social media has become an arena for producing opinion that is not always transparent. Astuti and Evelina specifically demonstrate the use of deepfakes in hoax narratives during the 2024 election campaign, while Rahman and Anggriawan frame the issue as an electoral crime that has not yet been clearly regulated. This article takes a more normative position by formulating a design for prohibiting political deepfakes within the election-campaign regime rather than merely mapping the risks of the technology.²

The urgency of this inquiry arises because deepfakes operate differently from ordinary textual hoaxes. Synthetic content can falsify a candidate's face, voice, gestures, or statements so that the public appears to witness an event that actually never occurred. In a short campaign period, clarification often arrives only after voters' perceptions have been shaped by repetition, emotion, and recommendation algorithms. Nurdin finds that disinformation and trolls in the 2024 election reinforced political polarization through provocative narratives that circulated continuously in digital spaces.³

The legal problem lies in the absence of a specific rule distinguishing political deepfakes from defamation, insult, hoaxes, or conventional black campaigning. The Election Law, the Electronic Information and Transactions Law, the Personal Data Protection Law, the KPU Regulation on Campaigning, and the Bawaslu Regulation on Campaign Supervision provide initial instruments but do not establish a category of violation for AI-based identity manipulation. Consequently, election supervisors may struggle to determine the type of violation, digital-evidence standards, handling

¹ Jürgen Habermas, *A New Structural Transformation of the Public Sphere and Deliberative Politics* (Cambridge: Polity Press, 2023), 7.

² Meily Astuti and Shane Evelina, "Penggunaan Deepfake Terkait Penyebaran Isu Hoaks Kampanye pada Pemilu 2024," *Jurnal Bawaslu DKI Jakarta* 9, no. 1 (2024): 1-14, <https://journal.bawaslu.go.id/index.php/JBK/article/view/441>.

³ Reskiyanti Nurdin, "Trolls, Disinformasi, dan Strategi Kampanye: Tantangan Demokrasi Digital dalam Pemilu 2024," *Electoral Governance: Jurnal Tata Kelola Pemilu Indonesia* 6, no. 1 (2025): 1-26, <https://journal.kpu.go.id/index.php/TKP/article/view/1350>.

deadlines, and the legal basis for ordering content removal. This operational gap may diminish the effectiveness of the principles of honesty and fairness in digital campaigning.⁴

This article addresses the need to regulate political deepfakes in Indonesian election campaigns. It focuses on three matters: the position of deepfakes in relation to the integrity of voter intent, weaknesses in current election law, and a proportionate prohibition model. The proposed model combines a ban on manipulative content, mandatory labeling of AI content, rapid clarification and removal mechanisms, and institutional coordination among the KPU, Bawaslu, Kominfo, and digital platforms. Under this design, the principles of honesty and fairness move from abstract normative principles to concrete controls on AI-based campaign manipulation.⁵

Existing Indonesian scholarship has largely mapped deepfakes as disinformation, cybercrime, or a technological risk. The underdeveloped question is narrower and doctrinal: what legal characteristics should trigger an election-law prohibition, and how can such a prohibition protect voter autonomy without suppressing parody, satire, legitimate AI-assisted production, or political criticism? This article therefore contributes a proportionality-based regulatory model that links the constitutional principles of honest and fair elections to content labeling, expedited remedies, evidentiary verification, and institutional coordination.

RESEARCH METHOD

This article adopts normative legal research using statutory, conceptual, case, and limited comparative approaches. Primary materials are read hierarchically and functionally: the 1945 Constitution supplies the constitutional standard; election legislation and KPU/Bawaslu regulations define campaign duties and supervisory authority; the Electronic Information and Transactions Law and Personal Data Protection Law are examined only to the extent that they address identity manipulation, unlawful electronic content, and personal-data processing. The case approach is used to test how existing categories operate when audiovisual identity is synthetically fabricated. The comparative materials are not transplanted as binding law, but are used as functional benchmarks for disclosure, timing, takedown, and freedom-of-expression safeguards. The analysis proceeds through four tests: legal object protected, deceptive character, electoral nexus and timing, and proportionality of the proposed remedy.

RESULTS AND DISCUSSION

The discussion proceeds through three layers of analysis. First, political deepfakes are examined as a threat to the integrity of campaign information and the authenticity of voter intent. Second, Indonesian positive law is tested to determine the capacity of election law, electronic-information law, and campaign-supervision instruments to address AI-based manipulation. Third, the article formulates a

⁴ Republik Indonesia, Undang-Undang Nomor 7 Tahun 2017 tentang Pemilihan Umum.

⁵ Atef Fahrudin and Syamsul Arif Billah, "Modeling Manajemen Informasi dan Verifikasi Berita Menjelang Pemilu dan Pemilihan 2024: Sebuah Studi Preliminary tentang Upaya Mencegah Black Campaign dan Hoax," *Electoral Governance: Jurnal Tata Kelola Pemilu Indonesia* 4, no. 2 (2023): 212-232, <https://doi.org/10.46874/tkp.v4i2.672>.

prohibition model that distinguishes manipulative content from lawful political expression.

Table 1. Comparison of Campaign-Information Violations

Form	Medium	Manipulative Effect	Legal Response
Textual hoax	Text, poster, screenshot	Misleads through a false narrative	Clarification, correction, the Electronic Information and Transactions Law, and specified election offenses
Black campaign	Speech, leaflet, post, advertisement	Attacks a candidate through defamation or insult	Campaign prohibitions and criminal or administrative enforcement
Political deepfake	Video, audio, image, synthetic avatar	Falsifies identity and events so that they appear authentic	A specific prohibition, AI labeling, digital forensics, and rapid takedown are required

Source: Analyzed from primary legal materials and literature on electoral disinformation.

Political Deepfakes and the Integrity of Voter Intent

Political deepfakes must be understood as an electoral-integrity problem because they attack the process through which voters form their will. In honest and fair elections, voters are entitled to assess contestants on the basis of information that can be substantiated. When a candidate's face or voice is manipulated to convey a false statement, voters do not merely receive incorrect information; they are induced to respond to a political object that never actually existed.⁶

The manipulative nature of deepfakes is evident in their capacity to imitate visual and auditory evidence. People generally trust audio or video recordings more readily than anonymous text because audiovisual media are perceived as traces of actual events. Chesney and Citron describe deepfakes as a challenge to democracy, privacy, and national security because the technology blurs the boundary between evidence and fabrication. In an election campaign, that ambiguity can transform political competition into an arena of identity falsification.⁷

Indonesian literature on election hoaxes shows that disinformation operates through rapid dissemination and difficulty of verification. Juditha emphasizes that interaction with hoaxes on social media grows through patterns of sharing, commenting, and repetition among users. Deepfakes add further destructive force to that pattern because fabricated content appears more convincing than a textual claim. The more realistic the manipulation of faces and voices, the heavier the burden of clarification placed on victims and election administrators.⁸

Digital campaigning does not take place in a neutral environment. Buzzers, trolls, anonymous accounts, and content-distribution networks can expand the reach of manipulative material in a coordinated manner. Arianto shows that political buzzers operate within digital politics by producing conversations, reinforcing

⁶ Pippa Norris, *Why Electoral Integrity Matters* (Cambridge: Cambridge University Press, 2014), 4.

⁷ Robert Chesney and Danielle Citron, "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security," *California Law Review* 107, no. 6 (2019): 1753-1820, <https://doi.org/10.15779/Z38RV0D15J>.

⁸ Christiany Juditha, "Interaksi Komunikasi Hoax di Media Sosial serta Antisipasinya," *Jurnal Pekommas* 3, no. 1 (2018): 31-44, <https://doi.org/10.30818/jpkm.2018.2030104>.

narratives, and attacking political opponents. A deepfake disseminated through such a network can create a dual effect: identity falsification at the content level and manipulative amplification at the distribution level.⁹

From the perspective of election law, the principal harm caused by a deepfake is not limited to a candidate's reputation. The more fundamental harm is injury to the integrity of voter intent. Voters may cast ballots because they believe a false recording, come to hate a candidate because of a statement never made, or refrain from voting because they consider the entire political process contaminated. At that point, deepfakes impair both the fairness of the competition and citizens' political autonomy.¹⁰

The distinction between deepfakes and satire or parody must be established at the outset. Satire and parody remain forms of political criticism insofar as they are not designed to deceive voters about a person's identity, statements, or actions. Iswara shows that satirical hoaxes concerning political issues can become digital attacks when audiences no longer recognize the humorous context and instead use the material for political harassment. A prohibition on deepfakes must therefore assess intent, the form of the content, the context of distribution, and its potential to mislead.¹¹

This assessment shows that political deepfakes cannot be addressed solely as a digital-literacy issue. Literacy remains necessary, but it is insufficient when false content is professionally produced, distributed at scale, and released close to election day. Elections require an institutional response faster than ordinary clarification processes. Speed becomes a legal consideration because delayed removal may allow truth to be overwhelmed by virality.

The most dangerous electoral deepfakes are not limited to long videos impersonating a candidate. A short audio clip imitating a candidate's voice may be used to issue false instructions to volunteers, make provocative statements to a particular group, or fabricate a confession. Synthetic images can be used to shape voters' moral perceptions, while short video clips may circulate through messaging applications without context. Because voters often receive political content in brief fragments, a deepfake need not be perfect to cause electoral harm.

The effect of deepfakes does not always take the form of an immediate change in electoral choice. It may instead create generalized doubt about all political information, leaving voters unable to distinguish an official clarification from a strategic denial. This phenomenon is an erosion of epistemic trust: factual truth loses corrective force because the public assumes that all evidence can be manipulated. In elections, such conditions benefit actors who deliberately disrupt the information environment rather than actors who compete through ideas and programs.

⁹ Bambang Arianto, "Kontestasi Buzzer Politik dalam Langgam Politik Digital," *MASSIVE: Jurnal Ilmu Komunikasi* 2, no. 2 (2022): 28-42, <https://doi.org/10.35842/massive.v2i2.77>.

¹⁰ Rofi Aulia Rahman and Rizaldy Anggriawan, "Deepfake and Electoral Crimes: Criminal Law Perspectives from Indonesia, India, Pakistan, and the U.S.," *Indonesian Comparative Law Review* 7, no. 2 (2025): 1-21, <https://doi.org/10.18196/iclr.v7i2.26337>.

¹¹ A. A. Iswara, "Hoaks Satire sebagai Cyberbullying dalam Konteks Isu Politik," *Jurnal Pekommas* 6, no. 2 (2021): 65-74, <https://doi.org/10.30818/jpkm.2021.2060207>.

A further problem is the unequal burden of proof. A deepfake victim must refute fabricated content, demonstrate that the recording is inauthentic, and persuade the public within a short period. Disseminators, meanwhile, may hide behind account anonymity, claims of satire, or the assertion that they merely forwarded another person's post. This imbalance demonstrates why election law requires a specific procedure rather than leaving resolution entirely to public debate.

In Indonesia, these risks are intensified because much political information circulates through closed messaging groups. Content shared in such spaces is more difficult to monitor than public posts. Deepfake material can move from one group to another without a clear trail to its source and later reappear as a screenshot, edited clip, or new narrative. A digital-campaign supervision model must account for this closed distribution pattern through rapid-reporting channels and cross-platform cooperation.

The integrity of voter intent does not mean that voters must be shielded from all negative information about candidates. Criticism, investigative journalism, adverse records, and vigorous debate remain part of democratic campaigning. What must be prohibited is the falsification of identities and events that causes voters to assess candidates on the basis of fraudulently constructed facts. This distinction is essential so that the principles of honest and fair elections are not used to restrict political freedom excessively.

Weaknesses in Indonesian Election-Law Norms

Law Number 7 of 2017 on General Elections prohibits campaign content involving insult, defamation, incitement, provocation, and material that disrupts public order. These prohibitions are important but do not recognize the distinctive character of AI-based identity manipulation. A deepfake may be defamatory, but it does not always take the form of a verbal allegation; it may instead be a false recording that leads the public to believe that a candidate performed an act. This particular method has not been formulated as a separate campaign violation.¹²

KPU Regulation Number 15 of 2023 governs campaigns through social media and campaign materials, while Bawaslu Regulation Number 11 of 2023 governs campaign supervision. Both instruments authorize election administrators to supervise digital spaces. Neither, however, establishes a special procedure for AI content, such as metadata requirements, synthetic-content labels, rapid complaint channels, verification deadlines, or technical coordination with platforms. Deepfakes operate on a much shorter timescale than ordinary violation-handling processes.¹³

The Electronic Information and Transactions Law may apply where the dissemination of a deepfake satisfies the elements of unlawful content, defamation, false news, or electronic-information manipulation. The Personal Data Protection Law is also relevant because faces, voices, and biometric data concern personal identity. The weakness is that neither law was designed as a specific instrument for enforcing election-campaign rules. Urgent electoral interests, including protection

¹² Republik Indonesia, Undang-Undang Nomor 7 Tahun 2017 tentang Pemilihan Umum.

¹³ Komisi Pemilihan Umum Republik Indonesia, Peraturan KPU Nomor 15 Tahun 2023 tentang Kampanye Pemilihan Umum.

during the campaign-silence period and restoration of voter perceptions, are therefore not always addressed through general cyberlaw mechanisms.¹⁴

Constitutional Court Decision Number 166/PUU-XXI/2023 provides an important point of entry because the Court interpreted an election contestant's public image as part of campaign material. If a candidate's public image is an object regulated in campaigning, AI-based falsification of that image should not be treated as an ordinary disturbance. Manipulating a candidate's face and voice directly affects the political identity through which voters recognize the contestant. The prohibition of deepfakes therefore finds a constitutional basis in the principles of honest and fair elections.¹⁵

Fahrudin and Billah propose an information-management and news-verification model for the 2024 elections to prevent black campaigns and hoaxes. Their proposal provides a useful foundation, but deepfakes require the additional layers of digital forensics, content-provenance evidence, and AI labeling. Febriansyah also emphasizes the KPU's role in countering hoaxes through open information, official content, and clarification. In relation to deepfakes, information openness must be developed into a rapid-response protocol connecting election administrators with digital platforms.¹⁶

A further weakness lies in the division of authority. The KPU regulates campaign procedures, Bawaslu supervises and handles violations, Kominfo has the capacity to terminate access to or otherwise address digital content, and platforms control the infrastructure through which content is distributed. Without a coordinating rule, each institution may act independently. In a viral deepfake case, fragmented authority results in a slow and inconsistent response.¹⁷

Overly general regulation also creates uncertainty for election contestants. Candidates, campaign teams, content creators, and volunteers lack a clear boundary between lawful uses of AI and prohibited manipulation. AI may legitimately be used for graphic design, translation, internal voice simulation, or transparent voter education. Sound law must not only prohibit conduct but also provide standards that prevent digital political innovation from becoming electoral deception.

The law also fails to distinguish among content creators, first disseminators, amplification actors, and parties that finance production. In digital campaigning, these actors are not always the same. An official team may not create the content but may benefit from its dissemination; volunteers may distribute material without instruction; and buzzers may operate through informal contracts. Without a clear mapping of legal subjects, enforcement will stop at the last disseminator rather than reach the actor who designed the manipulation.

The campaign-silence period is the most vulnerable point. A deepfake released one or two days before voting may influence perceptions when opportunities for

¹⁴ Republik Indonesia, Undang-Undang Nomor 1 Tahun 2024 tentang Perubahan Kedua atas Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik.

¹⁵ Mahkamah Konstitusi Republik Indonesia, Putusan Nomor 166/PUU-XXI/2023, 2024.

¹⁶ Atef Fahrudin and Syamsul Arif Billah, "Modeling Manajemen Informasi dan Verifikasi Berita Menjelang Pemilu dan Pemilihan 2024: Sebuah Studi Preliminary tentang Upaya Mencegah Black Campaign dan Hoax," *Electoral Governance: Jurnal Tata Kelola Pemilu Indonesia* 4, no. 2 (2023): 212-232, <https://doi.org/10.46874/tkp.v4i2.672>.

¹⁷ Badan Pengawas Pemilihan Umum Republik Indonesia, Peraturan Bawaslu Nomor 11 Tahun 2023 tentang Pengawasan Kampanye Pemilihan Umum.

formal clarification are severely limited. Indonesian election law prohibits campaigning during the silence period but provides no special procedure for a viral synthetic-content violation. The legal response must distinguish between urgent content removal and the examination of a violation, which still requires full proof.

Rules of evidence must also address the chain of custody for digital evidence. Post links, metadata, file hashes, screen recordings, expert testimony, and dissemination histories should be treated as an integrated evidentiary chain rather than as isolated items. Without such a procedure, proof of a deepfake can easily be contested because digital content can be re-edited, deleted, or moved to another platform. Bawaslu's technical guidelines should specify minimum standards for securing evidence from the moment a report is received.

Another weakness is the absence of substantial transparency obligations for election contestants that use AI to produce campaign materials. AI use is not inherently harmful: it may assist with visual design, translation, alternative text, or accessibility for voters with disabilities. When AI is used to imitate a person's face or voice, however, voters must be clearly informed so that they do not mistake the material for an authentic recording.

The supervision regime also does not require platforms to prioritize reports of manipulative electoral content. Digital platforms have their own moderation systems, but their pace does not always match the speed required during an election. If a Bawaslu report is treated like an ordinary user report, electoral harm may occur before the content is removed. A special protocol is therefore required during the campaign, the silence period, and election day.

The regulatory gap must not be filled by conferring overly broad powers without safeguards. A content-removal order must have a legal basis, a time limit, written reasons, and an objection mechanism. Due process remains necessary because political content often occupies a gray area among criticism, parody, propaganda, and manipulation. Without a right to object, deepfake enforcement could become a tool for silencing political opponents.

A Model for Prohibiting Political Deepfakes

A prohibition model for political deepfakes must begin with a precise definition. A political deepfake may be defined as audio, visual, or audiovisual content created, altered, or synthesized using artificial intelligence so that it depicts a person as saying, doing, or experiencing something untrue, where the content concerns an election contestant, election administrator, voter, or electoral issue. The definition covers not only video but also voice, images, avatars, and combinations of synthetic content. Its scope is deliberately tied to the electoral context so that it does not become a general prohibition on all digital expression.¹⁸

The prohibited element must be confined to manipulative and misleading content. This limitation is essential so that strong criticism, political memes, parody, satire, and creative works remain protected when they do not masquerade as authentic events. The rule may include an exception for content clearly labeled as parody, satire, education, or simulation and not disseminated with intent to deceive

¹⁸ Ruben Tolosana et al., "Deepfakes and Beyond: A Survey of Face Manipulation and Fake Detection," *Information Fusion* 64 (2020): 131-148, <https://doi.org/10.1016/j.inffus.2020.06.014>.

voters. Election law can thereby protect information integrity without turning campaign supervision into political censorship.¹⁹

Mandatory labeling is the second instrument. Any campaign content that uses AI to synthetically generate a candidate's voice, face, movements, or statements should carry a marker that can be readily seen or heard. The EU Artificial Intelligence Act establishes transparency requirements for synthetic content, including obligations designed to ensure that users know when specified content has been generated or manipulated by AI. Indonesia can adapt this principle in a KPU regulation or an amendment to the Election Law using language suited to the national electoral institutional structure.²⁰

The third instrument is a rapid clarification and removal mechanism. Bawaslu should have a priority channel for receiving reports of political deepfakes, requesting preliminary verification, ordering an interim clarification, and coordinating with Kominfo and platforms to remove manifestly manipulative content. California's AB 2655 provides an example of regulation targeting deceptive media in an electoral context and imposing specified obligations on platforms. Indonesia need not reproduce that model, but it can adopt its logic of short deadlines and mandatory platform response.²¹

The fourth instrument is a digital-evidence standard. Proof should not depend solely on admissions by the creator or the victim but should permit analysis of metadata, traces of the first upload, audiovisual forensics, expert testimony, and account-distribution patterns. Technical evidence should be combined with contextual evidence, including timing, accompanying narratives, political targets, and relationships with campaign teams. Ranka and colleagues emphasize that deepfake threats to electoral integrity involve interconnected processes of production, detection, and institutional response.²²

The fifth instrument is graduated sanctions. Minor violations may result in correction orders, mandatory clarification, or content removal. Serious violations that are coordinated, repeated, financed by an election contestant, or cause concrete electoral harm may attract administrative campaign sanctions and be referred for criminal proceedings where the elements of an offense are satisfied. Graduated sanctions avert both tolerance of grave manipulation and disproportionate criminalization of remediable technical errors.²³

This model requires stronger institutional capacity. The KPU should incorporate AI-labeling obligations into campaign rules; Bawaslu should establish a synthetic-content verification unit; Kominfo should provide a technical channel for

¹⁹ Maria Pawelec, "Deepfakes and Democracy (Theory): How Synthetic Audio-Visual Media for Disinformation and Hate Speech Threaten Core Democratic Functions," *Digital Society* 1 (2022): 1-37, <https://doi.org/10.1007/s44206-022-00010-6>.

²⁰ European Parliament and Council of the European Union, Regulation (EU) 2024/1689 Laying Down Harmonised Rules on Artificial Intelligence, 13 June 2024, Article 50.

²¹ State of California, Assembly Bill No. 2655, Elections: Deceptive Media in Advertisements, 2024.

²² Hriday Ranka et al., "Examining the Implications of Deepfakes for Election Integrity," arXiv, 2024, <https://arxiv.org/abs/2406.14290>.

²³ Khairul Fahmi, Feri Amsari, Busyra Azheri, and Muhammad Ichsan Kabullah, "Sistem Keadilan Pemilu dalam Penanganan Pelanggaran dan Sengketa Proses Pemilu Serentak 2019 di Sumatera Barat," *Jurnal Konstitusi* 17, no. 1 (2020): 1-26, <https://doi.org/10.31078/jk1711>.

handling electoral content; and platforms should retain dissemination-trace data during campaign periods. Putra and Arianto show that digital political dynamics before the 2024 election were inseparable from buzzers, coordinated narratives, and opinion formation on social media. A deepfake prohibition must therefore regulate content while also accounting for its distribution pattern.²⁴

The resulting rule may be formulated as follows: every person is prohibited from creating, disseminating, financing, or ordering the dissemination of AI-based campaign content that falsifies the identity, voice, face, statements, or actions of an election contestant, election administrator, or another party connected with an election where the content may mislead voters or impair campaign fairness. The rule should be accompanied by mandatory AI-content labeling, an exception for clearly identifiable parody, a rapid-removal procedure, and an objection mechanism. Under this design, election law is not anti-technology; it is anti-manipulation.

Table 2. Components of the Political Deepfake Prohibition Model

Component	Regulatory Content	Legal Function
Prohibition	AI content that falsifies electoral identity, voice, face, statements, or actions	Protects the integrity of voter choice
Labeling	A clear visual or audio marker on AI-generated campaign content	Prevents voters from mistaking synthetic content for an authentic recording
Rapid takedown	Initial clarification, verification, removal order, and right to object	Reduces viral harm without eliminating due process
Evidence	Metadata, file hashes, upload trails, forensic experts, and distribution patterns	Prevents enforcement from depending on unilateral claims
Graduated sanctions	Correction, removal, administrative sanctions, and criminal punishment for serious violations	Preserves proportionality between digital innovation and prevention of manipulation

Source: Author analysis based on election law, campaign supervision rules, and comparative AI transparency models.

Institutionally, the model can operate in three stages. The first is prevention, consisting of an obligation for election contestants to declare AI use in specified campaign materials and a prohibition on producing content that imitates another party's identity without consent. The second is rapid response, consisting of a reporting channel, preliminary verification, public clarification, and a request for content removal. The third is enforcement, consisting of an examination of the violation, identification of responsible legal subjects, and sanctions corresponding to the degree of culpability.

The prohibition should be enacted at the statutory level because it restricts freedom of expression and may carry sanctions. Technical rules can be prescribed in KPU and Bawaslu regulations, but the legal basis of the prohibition, its subjects and objects, exceptions, and types of sanction must be contained in the Election Law. Statutory placement provides stronger legitimacy and reduces disputes over the

²⁴ Ariandi Putra, "Peran Buzzer Politik dalam Dinamika Jelang Pemilu Tahun 2024," *SALAM: Jurnal Sosial dan Budaya Syar-i* 10, no. 4 (2023): 1143-1158, <https://doi.org/10.15408/sjsbs.v10i4.34076>.

authority of election administrators. Technical regulations would then explain procedures rather than create new prohibitions without a mandate.

The regulatory design must also account for the relationship between election law and criminal law. Not every political deepfake should attract criminal punishment because criminal law is an ultimum remedium. Content that is promptly corrected, causes no extensive harm, or is created without intent to deceive may be addressed through administrative sanctions. Criminal punishment is appropriate only where there is malicious intent, coordinated dissemination, political financing, serious harm, or large-scale use of a person's identity to deceive voters.

The labeling requirement should be simple enough for election contestants and content creators to follow. A label may state, 'This content was created or modified using artificial intelligence,' and appear at the beginning of a video, in a post description, or as an audio marker for voice content. An overly technical label will be ineffective because voters require information that is quick and easy to understand. The purpose is not to prohibit AI but to provide context so voters understand the nature of the content they consume.

A rapid-removal mechanism must be distinguished from a final ruling on a violation. At the preliminary stage, Bawaslu need only determine whether content appears to contain electoral-identity manipulation and may mislead voters. That assessment can justify a request for temporary restriction but should not immediately establish a person's guilt. A substantive examination would follow after complete evidence is collected, allowing electoral protection and defense rights to operate together.

The model also requires a public archive of clarified content. The purpose of the archive is not to recirculate deepfakes but to inform voters that specified content has been verified as manipulative. The KPU and Bawaslu could establish a joint clarification page containing a case summary, verification status, and a channel for further checking. Such transparency is important because removing content without explanation often gives rise to new conspiracy theories.

Ultimately, the prohibition of political deepfakes must be constructed as an electoral-integrity instrument rather than an anti-technology agenda. AI may continue to be used for voter education, accessibility, campaign design, and efficient political communication. What is prohibited is the use of AI to disguise a lie as authentic evidence. This boundary makes the model compatible with constitutional democracy: freedom of expression is respected, while manipulation that corrupts voter intent is not tolerated.

Comparative and Doctrinal Calibration

A Scopus-oriented framing should avoid assuming that every deepfake necessarily changes a vote. Experimental and communication research more cautiously indicates effects on uncertainty, trust, candidate evaluation, and polarization. The legal justification for intervention is therefore strongest when it is anchored in the integrity of the information environment and the authenticity of attributed political speech rather than in an unprovable claim of direct vote conversion.

Comparative scholarship also supports a graduated response. A workable model distinguishes (1) materially deceptive synthetic impersonation intended to be taken as authentic, (2) disclosed or readily recognizable parody and satire, (3)

benign AI-assisted editing, and (4) synthetic campaign advertising that is not necessarily false but still requires disclosure. This classification allows prohibitions to be narrowly tailored while preserving political expression.

The existing tables should be read as analytical devices rather than as independent empirical findings. Table 1 differentiates deepfakes from adjacent campaign-information violations by legal object, mode of deception, and regulatory consequence, while Table 2 converts that distinction into an operational prohibition model. Together, the tables show why a technology-specific response must remain content- and harm-sensitive.

CONCLUSION

Political deepfakes create a specific election-law problem when synthetic audio, images, or video materially falsify the identity, statement, or conduct of an electoral actor and are disseminated in circumstances reasonably capable of misleading voters. Indonesian law contains partial tools in election, electronic-information, and personal-data regimes, but those tools do not yet provide a single operational category with clear elements, evidence standards, response deadlines, and remedies. The proposed model therefore combines a narrowly defined prohibition on materially deceptive political impersonation, mandatory disclosure for other synthetic campaign content, expedited clarification and takedown, preservation of digital evidence, and coordinated authority among election supervisors, election administrators, the communications regulator, and platforms. The model should expressly exempt bona fide satire, parody, and clearly disclosed creative uses unless independently unlawful. Its central contribution is to translate the constitutional requirements of honest and fair elections into a proportionate governance framework for synthetic campaign media.

REFERENCES

- Akmaliah, Wahyudi. "Bukan Sekedar Penggaung (Buzzers): Media Sosial dan Transformasi Arena Politik." *Maarif* 13, no. 1 (2018): 9-25. <https://doi.org/10.47651/mrf.v13i1.9>.
- Anggraeni, Inez. "Deepfakes, Trust, and Electoral Integrity: The Emerging Threat to Indonesian Democracy." *Indonesian Discourse on Communication, Democracy, and Political Movements* 2, no. 4 (2025). <https://doi.org/10.65815/zrnx7164>.
- Arianto, Bambang. "Kontestasi Buzzer Politik dalam Langgam Politik Digital." *MASSIVE: Jurnal Ilmu Komunikasi* 2, no. 2 (2022): 28-42. <https://doi.org/10.35842/massive.v2i2.77>.
- Arianto, Bambang. "Salah Kaprah Ihwat Buzzer: Analisis Percakapan Warganet di Media Sosial." *JIP: Jurnal Ilmiah Ilmu Pemerintahan* 5, no. 1 (2020): 1-20. <https://doi.org/10.14710/jiip.v5i1.7287>.
- Carvajal, Laura, and Andrew Iliadis. "Deepfakes: A Preliminary Systematic Review of the Literature." *Selected Papers of Internet Research* (2020). <https://doi.org/10.5210/spir.v2020i0.11190>.
- Chesney, Robert, and Danielle Citron. "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security." *California Law Review* 107, no. 6 (2019): 1753-1820. <https://doi.org/10.15779/Z38RV0D15J>.
- Dobber, Tom, Nadia Metoui, Damian Trilling, Natali Helberger, and Claes de Vreese. "Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes?"

- International Journal of Press/Politics 26, no. 1 (2021). <https://doi.org/10.1177/1940161220944364>.
- Fahmi, Khairul, Feri Amsari, Busyra Azheri, and Muhammad Ichsan Kabullah. "Sistem Keadilan Pemilu dalam Penanganan Pelanggaran dan Sengketa Proses Pemilu Serentak 2019 di Sumatera Barat." *Jurnal Konstitusi* 17, no. 1 (2020): 1-26. <https://doi.org/10.31078/jk1711>.
- Fahrudin, Atef, and Syamsul Arif Billah. "Modeling Manajemen Informasi dan Verifikasi Berita Menjelang Pemilu dan Pemilihan 2024: Sebuah Studi Preliminary tentang Upaya Mencegah Black Campaign dan Hoax." *Electoral Governance: Jurnal Tata Kelola Pemilu Indonesia* 4, no. 2 (2023): 212-232. <https://doi.org/10.46874/tkp.v4i2.672>.
- Hidayat, Rayhan Naufaldi. "Penggunaan Buzzer Politik di Media Sosial pada Masa Kampanya Pemilihan Umum." *ADALAH* 4, no. 2 (2020): 29-38. <https://doi.org/10.15408/adalah.v4i2.15606>.
- Irvan, Rachmat, and Junet Hariyo Setiawan. "Criminalization Policy of Deepfake in Indonesian Elections." *Constitutional Law Society* 5, no. 1 (2026). <https://doi.org/10.36448/jcls.v5i1.125>.
- Iswara, A. A. "Hoaks Satire sebagai Cyberbullying dalam Konteks Isu Politik." *Jurnal Pekommas* 6, no. 2 (2021): 65-74. <https://doi.org/10.30818/jpkm.2021.2060207>.
- Juditha, Christiany. "Interaksi Komunikasi Hoax di Media Sosial serta Antisipasinya." *Jurnal Pekommas* 3, no. 1 (2018): 31-44. <https://doi.org/10.30818/jpkm.2018.2030104>.
- Kruschinski, Simon, and Fabio Votta. "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 (2026): 79-89. <https://doi.org/10.1007/s41358-026-00457-w>.
- Momeni, Mina. "Artificial Intelligence and Political Deepfakes: Shaping Citizen Perceptions Through Misinformation." *Journal of Creative Communications* 20, no. 1 (2025). <https://doi.org/10.1177/09732586241277335>.
- Neyazi, Taberez Ahmed, Tan Khai Ee, and Ozan Kuru. "Campaign Deepfakes and Affective Polarization: The Role of Artificial Intelligence in Campaigns in Shaping Voter Attitudes." *Social Science Computer Review* (2025). <https://doi.org/10.1177/08944393251362247>.
- Pawelec, Maria. "Deepfakes and Democracy (Theory): How Synthetic Audio-Visual Media for Disinformation and Hate Speech Threaten Core Democratic Functions." *Digital Society* 1 (2022): 1-37. <https://doi.org/10.1007/s44206-022-00010-6>.
- Putra, Ariandi. "Peran Buzzer Politik dalam Dinamika Jelang Pemilu Tahun 2024." *SALAM: Jurnal Sosial dan Budaya Syar-i* 10, no. 4 (2023): 1143-1158. <https://doi.org/10.15408/sjsbs.v10i4.34076>.
- Rahman, Rofi Aulia, and Rizaldy Anggriawan. "Deepfake and Electoral Crimes: Criminal Law Perspectives from Indonesia, India, Pakistan, and the U.S." *Indonesian Comparative Law Review* 7, no. 2 (2025): 1-21. <https://doi.org/10.18196/iclr.v7i2.26337>.

- Shukla, Anmol K. "AI-Generated Misinformation in the Election Year 2024." *Frontiers in Political Science* 6 (2024): 1-9. <https://doi.org/10.3389/fpos.2024.1451601>.
- Tolosana, Ruben, Ruben Vera-Rodriguez, Julian Fierrez, Aythami Morales, and Javier Ortega-Garcia. "Deepfakes and Beyond: A Survey of Face Manipulation and Fake Detection." *Information Fusion* 64 (2020): 131-148. <https://doi.org/10.1016/j.inffus.2020.06.014>.
- Vaccari, Cristian, and Andrew Chadwick. "Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News." *Social Media + Society* 6, no. 1 (2020). <https://doi.org/10.1177/2056305120903408>.
- Zhang, Bingbing, Sang Jung Kim, and Alex Scott. "Immunizing the Public Against AI-Generated Disinformation: Testing the Effects of Inoculation Mode and Issue Attitude on Inoculation Likelihood of Political Deepfakes." *Journalism & Mass Communication Quarterly* 102, no. 4 (2025). <https://doi.org/10.1177/10776990251357949>.
- Łabuz, Mateusz, and Christopher Nehring. "On the Way to Deep Fake Democracy? Deep Fakes in Election Campaigns in 2023." *European Political Science* 23 (2024): 454-473. <https://doi.org/10.1057/s41304-024-00482-9>.
- Łabuz, Mateusz. "Political Advertising with the Use of Deep Fakes—Can Deliberative Democracy Really Benefit from This Strategy?" *Politics & Policy* (2025). <https://doi.org/10.1111/polp.70035>.

Primary Legal Materials

- Badan Pengawas Pemilihan Umum Republik Indonesia. Peraturan Bawaslu Nomor 11 Tahun 2023 tentang Pengawasan Kampanye Pemilihan Umum.
- European Parliament and Council of the European Union. Regulation (EU) 2024/1689 Laying Down Harmonised Rules on Artificial Intelligence. 13 June 2024.
- Komisi Pemilihan Umum Republik Indonesia. Peraturan KPU Nomor 15 Tahun 2023 tentang Kampanye Pemilihan Umum.
- Mahkamah Konstitusi Republik Indonesia. Putusan Nomor 166/PUU-XXI/2023. 2024.
- Republik Indonesia. Undang-Undang Dasar Negara Republik Indonesia Tahun 1945.
- Republik Indonesia. Undang-Undang Nomor 1 Tahun 2024 tentang Perubahan Kedua atas Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik.
- Republik Indonesia. Undang-Undang Nomor 27 Tahun 2022 tentang Perlindungan Data Pribadi.
- Republik Indonesia. Undang-Undang Nomor 7 Tahun 2017 tentang Pemilihan Umum.