Indonesian Political Transformation
ISSN: XXXX-XXXX (Print) XXXX-XXXX (Online)
Volume 1, Number 2 (2025): 229-254
Available online since October 30, 2025

Regulating Political Deepfakes in Indonesia: Democratic Governance and Electoral Integrity in Southeast Asia

Mengatur Deepfake Politik di Indonesia: Tata Kelola Demokrasi dan Integritas Pemilu di Asia Tenggara

Yasmin Aurelia Hapsari, *Universitas Muhammadiyah Tangerang*

Abstract

The rapid diffusion of generative artificial intelligence has introduced unprecedented challenges to democratic governance, particularly through the emergence of political deepfakes capable of manipulating public opinion and undermining electoral integrity. While scholarship on AI-generated disinformation has expanded rapidly in advanced democracies, limited attention has been devoted to emerging democracies where digital literacy, regulatory capacity, and electoral institutions remain uneven. This study examines Indonesia's regulatory response to political deepfakes within the broader context of democratic governance in Southeast Asia. Employing a qualitative case study approach, the research analyzes Indonesia's legal framework governing digital information, electoral regulations, AI-related policies, and institutional responses during the 2024 general election. Documentary analysis is complemented by interviews with policymakers, election officials, civil society organizations, and digital platform experts to identify governance gaps and regulatory challenges. The findings are expected to demonstrate that existing legal instruments addressing misinformation are insufficient to

Publisher: Artes Libres Institutum
Available online **at** https://journals.arteslibres.org/index.php/politics

cope with AI-generated synthetic media because they were designed for conventional forms of digital manipulation. Furthermore, fragmented institutional coordination and limited technological capacity constrain effective enforcement. The article argues that regulating political deepfakes requires an integrated governance model combining adaptive legislation, platform accountability, AI detection mechanisms, public digital literacy, and regional cooperation among Southeast Asian democracies. By positioning Indonesia as an important case within the Global South, this study contributes to emerging debates on AI governance, electoral integrity, and democratic resilience. The findings offer policy recommendations for balancing freedom of expression with the protection of democratic processes in increasingly AI-mediated political environments.

Keywords: *Political Deepfakes, Artificial Intelligence Governance, Electoral Integrity, Indonesia, Democratic Resilience*

Abstrak

Penyebaran pesat kecerdasan buatan generatif telah menghadirkan tantangan yang belum pernah terjadi sebelumnya bagi tata kelola demokrasi, khususnya melalui munculnya deepfake politik yang mampu memanipulasi opini publik dan merusak integritas pemilu. Meskipun kajian tentang disinformasi yang dihasilkan AI telah berkembang pesat di negara-negara demokrasi maju, perhatian yang terbatas telah diberikan kepada negara-negara demokrasi berkembang di mana literasi digital, kapasitas regulasi, dan lembaga pemilu masih belum merata. Studi ini meneliti respons regulasi Indonesia terhadap deepfake politik dalam konteks yang lebih luas dari tata kelola demokrasi di Asia Tenggara. Dengan menggunakan pendekatan studi kasus kualitatif, penelitian ini menganalisis kerangka hukum Indonesia yang mengatur informasi digital, peraturan pemilu, kebijakan terkait AI, dan respons kelembagaan selama pemilihan umum 2024. Analisis dokumen dilengkapi dengan wawancara dengan para pembuat kebijakan, pejabat pemilu, organisasi masyarakat sipil, dan pakar platform digital untuk mengidentifikasi kesenjangan tata kelola dan tantangan regulasi. Temuan diharapkan menunjukkan bahwa instrumen hukum yang ada untuk mengatasi misinformasi tidak cukup untuk mengatasi media sintetis yang dihasilkan AI karena

dirancang untuk bentuk manipulasi digital konvensional. Lebih lanjut, koordinasi kelembagaan yang terfragmentasi dan kapasitas teknologi yang terbatas membatasi penegakan hukum yang efektif. Artikel ini berpendapat bahwa pengaturan deepfake politik membutuhkan model tata kelola terintegrasi yang menggabungkan legislasi adaptif, akuntabilitas platform, mekanisme deteksi AI, literasi digital publik, dan kerja sama regional di antara negara-negara demokrasi Asia Tenggara. Dengan memposisikan Indonesia sebagai studi kasus penting di Global South, penelitian ini berkontribusi pada perdebatan yang muncul tentang tata kelola AI, integritas pemilu, dan ketahanan demokrasi. Temuan-temuan tersebut menawarkan rekomendasi kebijakan untuk menyeimbangkan kebebasan berekspresi dengan perlindungan proses demokrasi dalam lingkungan politik yang semakin dimediasi oleh AI.

Kata kunci: *Deepfake Politik, Tata Kelola Kecerdasan Buatan, Integritas Pemilu, Indonesia, Ketahanan Demokrasi*

I. INTRODUCTION

The rapid advancement of generative artificial intelligence (AI) has fundamentally transformed the production, dissemination, and consumption of digital information. Among the most consequential developments is the emergence of deepfakes—AI-generated synthetic media capable of producing highly realistic audio, images, and videos that are often indistinguishable from authentic content. While generative AI offers significant opportunities for innovation across education, healthcare, and creative industries, it also introduces unprecedented risks to democratic governance by facilitating sophisticated forms of political disinformation, identity manipulation, and information warfare (Chesney and Citron 2019; Westerlund 2019). Unlike conventional misinformation, political deepfakes possess a unique capacity to fabricate persuasive visual evidence, thereby undermining public trust in information ecosystems and eroding confidence in democratic institutions.

Growing concern over political deepfakes has prompted governments, technology companies, and international organizations to reconsider existing regulatory approaches to digital information. Democratic societies increasingly recognize that traditional legal frameworks governing misinformation were largely designed for human-generated content and therefore struggle to address AI-generated synthetic media capable of being produced rapidly, anonymously, and at minimal cost (UNESCO 2023). Moreover, advances in open-source generative AI models have significantly lowered the technological barriers to creating convincing manipulated political content, enabling malicious actors to interfere with elections, damage political reputations, and amplify polarization through coordinated influence operations (OECD 2024). Consequently, regulating political deepfakes has become one of the most pressing challenges in contemporary AI governance.

The democratic implications of political deepfakes extend beyond the dissemination of false information. Existing scholarship suggests that synthetic media threatens the integrity of electoral processes by distorting informed political choice, weakening public confidence in electoral institutions, and creating what has been described as the "liar's dividend," whereby genuine evidence may be dismissed as fabricated because deepfake technology has become widely recognized (Chesney and Citron 2019). Such dynamics undermine one of democracy's fundamental requirements: the availability of reliable information upon which citizens can make informed electoral decisions. Consequently, debates surrounding political deepfakes increasingly intersect with broader discussions concerning democratic resilience, platform governance, algorithmic accountability, and the regulation of emerging artificial intelligence technologies (Gorwa 2019; UNESCO 2023).

These challenges are particularly significant in emerging democracies, where institutional capacities, digital literacy, and regulatory frameworks often remain uneven. While advanced

democracies such as the European Union and the United States have accelerated efforts to establish AI governance frameworks—including the European Union Artificial Intelligence Act and voluntary commitments among major technology companies—many countries in the Global South continue to rely upon legal instruments originally designed to regulate conventional online misinformation rather than AI-generated synthetic content (OECD 2024). This regulatory gap creates vulnerabilities during electoral processes, especially in highly digitalized societies where social media platforms serve as primary sources of political information.

Indonesia represents an especially important case for examining these governance challenges. As the world's third-largest democracy and one of Southeast Asia's largest digital economies, Indonesia has experienced rapid growth in internet penetration, social media use, and digital political campaigning. Digital platforms have become central arenas for political communication, electoral mobilization, and public debate, particularly during national and local elections. The 2024 General Election further demonstrated the increasing role of AI-assisted political communication, including the use of generative AI for campaign materials, political messaging, and public engagement. Although there is limited publicly verified evidence that large-scale malicious political deepfakes decisively influenced electoral outcomes, the widespread availability of generative AI technologies has heightened concerns regarding future electoral manipulation and the preparedness of existing regulatory institutions to address emerging technological risks.

Indonesia's current legal framework governing digital information primarily consists of the Electronic Information and Transactions (ITE) Law, Government Regulation No. 71 of 2019 concerning Electronic Systems and Transactions, electoral legislation administered by the General Elections Commission (KPU) and the Election Supervisory Agency (Bawaslu), and broader provisions relating to cybercrime and online misinformation. These instruments provide mechanisms for

addressing unlawful online content and digital misconduct; however, they were largely developed before the widespread diffusion of generative AI. As a result, existing regulations provide limited guidance regarding AI-generated synthetic media, the allocation of responsibility among platform providers and AI developers, standards for content authentication, or institutional procedures for responding rapidly to politically sensitive deepfake incidents during election periods.

The emergence of political deepfakes therefore exposes important governance challenges extending beyond legal regulation alone. Effective responses require institutional coordination among election management bodies, technology regulators, law enforcement agencies, digital platforms, civil society organizations, and AI developers. At the same time, regulatory interventions must remain consistent with democratic principles, particularly the protection of freedom of expression, political participation, and open public debate. Excessively restrictive regulation risks suppressing legitimate political speech and innovation, whereas insufficient regulation may enable increasingly sophisticated forms of electoral manipulation. This governance dilemma illustrates the broader challenge of balancing technological innovation with democratic accountability in rapidly evolving digital environments.

Despite the growing international literature on AI governance and synthetic media, existing scholarship remains heavily concentrated on advanced democracies, particularly the United States and Europe (Chesney and Citron 2019; UNESCO 2023). Comparatively fewer studies investigate how emerging democracies in the Global South confront the institutional challenges posed by political deepfakes. Furthermore, much of the existing research focuses on the technological capabilities of deepfake detection or the ethical implications of generative AI, while comparatively limited attention has been devoted to institutional governance, electoral regulation, and cross-sectoral coordination within democratic political systems. Consequently,

important theoretical and empirical gaps remain regarding how countries with developing regulatory institutions can design adaptive governance frameworks capable of responding effectively to rapidly evolving AI technologies.

This article addresses these gaps by examining Indonesia's regulatory response to political deepfakes within the broader context of democratic governance and electoral integrity in Southeast Asia. Rather than treating political deepfakes solely as a technological problem, the study conceptualizes them as a governance challenge requiring coordinated institutional responses across legal, technological, political, and societal domains. Employing a qualitative case study approach, the research analyzes Indonesia's legal framework governing digital information, electoral regulations, AI-related policy developments, and institutional responses during the 2024 General Election. The study argues that existing regulatory instruments remain insufficient because they were primarily designed to address conventional forms of digital misinformation rather than AI-generated synthetic media. Accordingly, it proposes an integrated governance model emphasizing adaptive legislation, platform accountability, AI detection mechanisms, digital literacy, and regional cooperation as complementary pillars for strengthening democratic resilience.

This study contributes to the literature in three important ways. First, it expands comparative scholarship on AI governance by providing empirical evidence from Indonesia, an underrepresented case within Global South research on political deepfakes. Second, it bridges debates on artificial intelligence governance and electoral integrity by demonstrating that effective regulation requires institutional coordination rather than isolated legal reforms. Third, it develops a governance-oriented analytical framework that may inform policy development across Southeast Asian democracies confronting similar challenges associated with the rapid diffusion of generative artificial intelligence. In doing so, the article contributes to broader discussions concerning how democratic societies can preserve electoral integrity while

safeguarding freedom of expression in increasingly AI-mediated political environments.

II. THE EMERGING THREAT OF POLITICAL DEEPFAKES TO ELECTORAL INTEGRITY IN INDONESIA

The rapid advancement of generative artificial intelligence (AI) has transformed political communication by fundamentally changing how political messages are created, disseminated, and consumed. Recent breakthroughs in large language models, image synthesis, voice cloning, and AI-generated video have significantly reduced the technical expertise and financial resources previously required to manipulate digital content. Today, sophisticated synthetic media can be produced using publicly available AI tools, enabling individuals and organizations to generate convincing political speeches, fabricated interviews, or manipulated campaign videos within minutes. While these technological developments have enhanced creativity and digital innovation, they have simultaneously increased the capacity for large-scale information manipulation capable of influencing democratic processes (Chesney and Citron 2019; Westerlund 2019).

Unlike conventional forms of digital misinformation, political deepfakes represent a qualitatively different category of electoral manipulation because they exploit the persuasive power of audiovisual authenticity. Human cognition generally assigns greater credibility to visual and auditory evidence than to textual information, making fabricated videos or audio recordings substantially more convincing than traditional false news or manipulated photographs. Consequently, political deepfakes possess an unprecedented capacity to distort public perception by presenting fabricated events as authentic evidence, thereby blurring the distinction between truth and fiction within increasingly complex digital information environments (Öhman 2020; Diakopoulos and Johnson 2021).

The democratic risks associated with political deepfakes extend beyond the immediate dissemination of false information. Chesney and Citron (2019) argue that the proliferation of synthetic media creates what they describe as the "liar's dividend," whereby the existence of convincing deepfake technologies allows political actors to dismiss authentic evidence as fabricated whenever it becomes politically inconvenient. Rather than merely facilitating deception, deepfakes undermine the epistemic foundations of democracy by weakening citizens' confidence in the authenticity of digital evidence itself. Subsequent research has demonstrated that the liar's dividend increases opportunities for political denial, reduces public trust in verified information, and complicates democratic accountability by allowing both fabricated and authentic media to become equally contestable.

This erosion of informational trust has become one of the defining governance challenges of the AI era. UNESCO argues that societies are approaching a "synthetic reality" in which citizens may no longer reliably distinguish authentic digital content from AI-generated fabrications without technological assistance. Because generative AI systems continue to improve more rapidly than detection technologies, regulators increasingly confront an asymmetric technological environment in which the production of synthetic media advances faster than institutional capacities for verification and enforcement. These developments indicate that deepfakes should no longer be understood merely as technological innovations but as systemic risks affecting democratic deliberation, electoral legitimacy, and public trust.

Political deepfakes present particularly significant risks during electoral periods because democratic elections depend upon citizens' ability to make informed political choices based on credible information. Electoral campaigns have become increasingly digitalized, with political parties, candidates, media organizations, and citizens relying extensively on social media platforms to communicate policy positions, mobilize supporters, and debate public issues. While digital platforms have expanded democratic

participation, they have also accelerated the circulation of emotionally charged and highly personalized political content that frequently prioritizes engagement over factual accuracy (Lazer et al. 2018; Bennett and Livingston 2018). Within such environments, AI-generated synthetic media can spread rapidly before verification mechanisms become effective, amplifying uncertainty precisely when voters require reliable information.

Indonesia provides an especially important context for examining these challenges. As one of the world's largest electoral democracies, Indonesia has experienced remarkable digital transformation over the past decade. Internet penetration has expanded rapidly, while platforms such as TikTok, Instagram, Facebook, YouTube, and WhatsApp have become primary channels of political communication, campaign mobilization, and voter engagement. During the 2024 General Election, political actors increasingly experimented with AI-assisted campaign materials, including AI-generated visual content, campaign advertisements, and digital branding strategies. Although publicly verified cases of malicious political deepfakes remained relatively limited, the growing accessibility of generative AI substantially increased concerns regarding future electoral vulnerabilities, particularly because synthetic media can be created and disseminated faster than existing regulatory institutions are capable of responding.

These concerns are consistent with broader international developments. Recent elections across multiple democratic jurisdictions have demonstrated that deepfakes are no longer hypothetical threats but practical instruments capable of disrupting political communication. AI-generated political audio circulated shortly before elections in Slovakia, fabricated robocalls imitating political leaders emerged in the United States, manipulated campaign materials appeared during electoral contests in South Asia, and synthetic political videos have increasingly been documented across diverse political systems. Collectively, these incidents illustrate that electoral manipulation through AI-

generated content has become a global governance challenge rather than a country-specific phenomenon.

The Indonesian case nevertheless presents distinctive institutional vulnerabilities. First, the country's high reliance on social media for political information increases the speed at which synthetic content may circulate across digital communities. Second, digital literacy levels remain uneven across regions and demographic groups, limiting citizens' capacity to distinguish authentic political communication from AI-generated manipulation. Third, electoral management institutions were primarily designed to regulate conventional campaign practices rather than rapidly evolving AI-generated content. Finally, Indonesia's existing legal framework governing misinformation predates the widespread adoption of generative AI, creating important regulatory ambiguities concerning responsibility, detection, verification, and enforcement.

These institutional characteristics suggest that political deepfakes should not simply be interpreted as another form of online misinformation. Rather, they represent an emerging category of AI-enabled electoral risk requiring governance approaches that integrate technological capacity, regulatory adaptation, democratic oversight, and institutional coordination. Existing misinformation policies largely assume that false information originates from identifiable human actors and can be evaluated through conventional fact-checking processes. Political deepfakes fundamentally challenge these assumptions because synthetic media increasingly combines algorithmic content generation, cross-platform dissemination, anonymous production, and transnational technological infrastructures.

From a governance perspective, therefore, the principal challenge is not solely detecting manipulated content but preserving democratic trust within increasingly AI-mediated information ecosystems. OECD (2024) emphasizes that strengthening information integrity has become a central requirement for democratic resilience, requiring governments to

combine regulatory interventions with platform accountability, public education, technological innovation, and institutional cooperation. Similarly, UNESCO highlights that sustainable responses to synthetic media cannot rely exclusively on technological detection because the pace of AI development continually outstrips purely technical countermeasures. Instead, democratic societies require governance systems capable of simultaneously protecting freedom of expression, encouraging technological innovation, and safeguarding electoral integrity against increasingly sophisticated forms of manipulation.

Accordingly, Indonesia's experience illustrates a broader transformation in contemporary electoral governance. The central issue is no longer whether misinformation exists within democratic societies but whether existing institutions possess sufficient adaptive capacity to regulate AI-generated political communication without undermining constitutional freedoms. Understanding this institutional challenge requires moving beyond technological explanations toward a governance-oriented analysis examining the adequacy of legal frameworks, institutional coordination, and regulatory preparedness. The following section therefore evaluates Indonesia's existing legal and institutional responses to political deepfakes in order to assess whether current governance arrangements are capable of protecting electoral integrity in an increasingly AI-mediated political environment.

III. REGULATORY READINESS: ASSESSING INDONESIA'S LEGAL AND INSTITUTIONAL RESPONSE TO POLITICAL DEEPFAKES

The emergence of political deepfakes challenges existing regulatory frameworks because most digital governance systems were originally developed to address conventional forms of misinformation rather than AI-generated synthetic media. Traditional approaches to information regulation generally assume that harmful content is created by identifiable human actors,

disseminated through recognizable channels, and can be addressed through established mechanisms of removal, accountability, or legal sanction. However, generative AI disrupts these assumptions by enabling automated content production, anonymous creation, rapid dissemination, and cross-border distribution. As a result, effective governance of political deepfakes requires regulatory systems capable of adapting to technological uncertainty while maintaining democratic safeguards (Gasser and Almeida 2017; Floridi et al. 2018).

Indonesia's existing regulatory framework for digital information governance has developed primarily through the Electronic Information and Transactions (ITE) Law, Government Regulation No. 71 of 2019 concerning Electronic Systems and Transactions, Ministerial Regulation No. 5 of 2020 concerning Private Electronic System Operators, and electoral regulations administered by the General Elections Commission (KPU) and the Election Supervisory Agency (Bawaslu). These instruments provide important legal foundations for addressing online harms, including illegal content, cybercrime, electronic fraud, and violations of electoral regulations. However, their regulatory assumptions were established before the widespread diffusion of generative AI technologies. Consequently, while existing frameworks provide mechanisms for responding to digital misconduct, they offer limited conceptual and operational guidance for addressing AI-generated political manipulation.

The limitation of Indonesia's current legal framework is particularly evident in relation to the definition and classification of synthetic media. Existing digital regulations generally focus on the legality of information content rather than the technological process through which such content is produced. This distinction becomes increasingly significant in the context of deepfakes because the central governance challenge is not only whether a particular message is false, but whether the process of creating and distributing synthetic media creates unique risks for democratic decision-making. A fabricated political video generated by AI may

involve multiple actors, including AI developers, platform providers, content creators, and distributors, making traditional models of individual responsibility increasingly difficult to apply.

This regulatory challenge reflects a broader issue identified within AI governance scholarship: technological innovation often develops at a faster pace than legal adaptation. Gasser and Almeida (2017) argue that AI governance requires regulatory approaches capable of managing uncertainty, technological complexity, and distributed responsibility rather than relying solely on traditional command-and-control regulation. Similarly, Floridi et al. (2018) emphasize that AI governance should be based on principles of transparency, accountability, human oversight, and societal benefit rather than narrow technological control. From this perspective, Indonesia's challenge is not merely the absence of AI-specific legislation but the need to redesign governance mechanisms capable of responding to continuously evolving technological capabilities.

A second challenge concerns institutional fragmentation. Political deepfake governance involves multiple actors with overlapping but sometimes unclear responsibilities, including the Ministry of Communication and Digital Affairs, election management institutions, law enforcement agencies, digital platforms, civil society organizations, and technology experts. Although each institution possesses relevant functions, the absence of a dedicated coordination mechanism creates potential gaps in detection, verification, response time, and public communication. During electoral periods, such fragmentation becomes particularly problematic because the impact of synthetic political content is highly time-sensitive. A manipulated video released shortly before voting may influence public perception before institutions have sufficient opportunity to investigate and respond.

The problem of institutional coordination is not unique to Indonesia. Research on digital governance demonstrates that contemporary technology-related risks increasingly require collaborative governance models because no single institution

possesses sufficient authority, expertise, or resources to manage complex digital ecosystems independently (Ansell and Gash 2008; Kooiman 2003). In the context of political deepfakes, government agencies require cooperation with technology companies that control detection systems and platform infrastructure, while civil society organizations contribute independent monitoring and public awareness capabilities. Therefore, effective governance depends not only on legal authority but also on institutional coordination and information-sharing mechanisms.

Platform accountability represents another critical dimension of Indonesia's regulatory preparedness. Social media companies occupy a central position in the governance of political deepfakes because they determine how synthetic content is detected, labeled, promoted, restricted, or removed. Platforms such as Meta, Google, TikTok, and YouTube have introduced various policies concerning AI-generated content, including labeling initiatives, manipulated media restrictions, and election integrity programs. However, these approaches remain largely based on voluntary corporate governance rather than binding national regulatory obligations. This creates important questions regarding transparency, consistency, and accountability.

Platform governance scholars argue that digital platforms have become powerful regulatory actors because their content moderation decisions directly influence public discourse and political participation (Gillespie 2018; Gorwa 2019). However, private governance mechanisms may produce accountability challenges because platforms operate according to complex combinations of commercial interests, community standards, government pressure, and global regulatory obligations. In the context of political deepfakes, reliance exclusively on platform self-regulation may be insufficient because companies may lack incentives, technological capacity, or contextual understanding necessary to evaluate politically sensitive synthetic content within specific national environments.

The European Union's approach provides an important comparative reference. Through the Digital Services Act (DSA) and the strengthened Code of Practice on Disinformation, the European regulatory model has moved toward greater platform transparency, risk assessment obligations, and accountability requirements for very large online platforms. While Indonesia's institutional and political context differs significantly from the European Union, this comparison demonstrates the importance of moving beyond reactive content removal toward proactive risk management and transparency-based governance. AI-generated political content requires systems that can identify risks before they produce significant democratic harm.

Another significant challenge concerns technological capacity. Effective regulation of political deepfakes requires access to reliable detection technologies, digital forensic expertise, and rapid verification mechanisms. However, many emerging democracies face limitations in technological infrastructure, financial resources, and specialized human capacity. Indonesia's regulatory institutions have made progress in digital governance, but the development of AI-specific monitoring capabilities remains a significant challenge. Without adequate technological capacity, legal frameworks may exist formally but remain difficult to implement effectively.

This situation reflects what scholars describe as the implementation gap in technology governance, where regulatory ambitions exceed institutional capabilities (Bannister and Connolly 2014). Strong regulations without sufficient enforcement capacity may create symbolic compliance rather than meaningful governance outcomes. In the case of political deepfakes, delayed detection and response mechanisms can significantly reduce regulatory effectiveness because synthetic content can achieve widespread circulation within a short period. Therefore, regulatory readiness must be evaluated not only through legal provisions but also through institutional resources, technological capabilities, and operational coordination.

Nevertheless, Indonesia's democratic context provides important opportunities for developing a more balanced governance approach. Unlike authoritarian systems where digital regulation may primarily function as a mechanism of political control, Indonesia possesses democratic institutions capable of providing oversight and contestation. Electoral institutions, judicial mechanisms, civil society organizations, journalists, and academic communities continue to play important roles in monitoring digital political developments. These democratic resources create opportunities for developing governance frameworks that address AI-related risks without undermining freedom of expression and political participation.

The analysis suggests that Indonesia's primary challenge is therefore not the absence of regulatory concern but the mismatch between rapidly evolving AI technologies and existing institutional arrangements. Current regulations provide partial tools for addressing harmful digital content, but political deepfakes require a more adaptive governance model capable of integrating legal reform, technological innovation, institutional coordination, and public participation. Without such adaptation, Indonesia risks relying on outdated regulatory mechanisms that are unable to effectively address increasingly sophisticated forms of electoral manipulation.

Therefore, Indonesia's regulatory readiness should be understood as an evolving process rather than a fixed condition. The country has developed important foundations for digital governance, but political deepfakes require a transition from conventional misinformation regulation toward comprehensive AI governance. This transition requires strengthening platform accountability, developing AI detection infrastructure, improving institutional coordination, and establishing clearer legal standards for synthetic political content. The following section develops an integrated governance framework that explains how emerging democracies can balance electoral protection, technological

innovation, and democratic freedoms in responding to AI-generated political manipulation.

IV. TOWARD AN INTEGRATED POLITICAL DEEPFAKE GOVERNANCE FRAMEWORK FOR SOUTHEAST ASIAN DEMOCRACIES

The analysis of Indonesia's regulatory preparedness demonstrates that political deepfakes represent a governance challenge that cannot be addressed through conventional approaches to misinformation regulation. Unlike traditional forms of false information, AI-generated synthetic media introduces a complex combination of technological uncertainty, distributed responsibility, and accelerated dissemination. Consequently, effective governance requires more than the expansion of legal prohibitions or content removal mechanisms. It requires an integrated institutional framework capable of coordinating regulatory, technological, societal, and regional responses while maintaining democratic principles.

This study therefore proposes an Integrated Political Deepfake Governance Framework as a conceptual approach for understanding how emerging democracies can respond to AI-generated electoral manipulation. The framework departs from purely technological or security-oriented approaches by positioning political deepfakes as a multidimensional governance problem involving law, institutions, platforms, citizens, and international cooperation. Rather than conceptualizing deepfake regulation as a question of restricting digital content, this framework emphasizes the creation of resilient democratic ecosystems capable of identifying, responding to, and mitigating AI-enabled information threats.

The first dimension of this framework is adaptive regulatory architecture. Existing regulatory approaches often struggle with emerging technologies because legal systems typically develop through slower institutional processes compared with rapid

technological innovation. Political deepfakes illustrate this regulatory mismatch because AI capabilities evolve continuously, while legislation frequently relies on fixed definitions and established categories of harm. Therefore, effective AI governance requires regulatory systems that are sufficiently flexible to accommodate technological change while maintaining legal certainty and democratic accountability.

Adaptive regulation does not imply unlimited governmental discretion. Instead, it requires establishing principles-based regulatory approaches that define clear responsibilities while allowing institutions to adjust implementation mechanisms as technologies evolve. The European Union's AI governance approach demonstrates this movement by adopting risk-based regulation rather than attempting to prohibit specific technologies entirely. Similarly, OECD AI Principles emphasize that trustworthy AI governance requires transparency, accountability, human oversight, and risk management throughout the AI lifecycle (OECD 2019). For emerging democracies such as Indonesia, the challenge is therefore to develop regulatory mechanisms that are flexible enough to address future AI developments while preventing arbitrary restrictions on political expression.

Within the Indonesian context, adaptive regulatory architecture requires revisiting existing digital information regulations to incorporate AI-specific considerations. Future regulatory reforms should clarify several issues, including the definition of political deepfakes, disclosure requirements for AI-generated political content, responsibilities of content creators and platforms, and procedures for responding to election-related synthetic media incidents. Without such adaptation, existing legal instruments may remain reactive rather than preventive, addressing harms only after manipulated content has already influenced public opinion.

The second dimension concerns platform accountability and technological governance. Digital platforms represent critical actors in managing political deepfakes because they control the

infrastructures through which synthetic media is produced, distributed, and amplified. Unlike traditional media organizations, platforms operate through algorithmic systems that determine visibility, engagement, and information circulation. Therefore, democratic governance increasingly depends on the accountability practices of private technology companies.

Platform accountability requires moving beyond voluntary commitments toward more structured governance mechanisms. These mechanisms may include mandatory transparency reporting, disclosure requirements for AI-generated political content, cooperation with independent fact-checking organizations, development of detection technologies, and clear procedures for responding to election-related manipulation. Gillespie (2018) argues that platforms have become central governance institutions because their moderation decisions shape the boundaries of public discourse. Consequently, their responsibilities should be evaluated not merely through commercial standards but also through democratic accountability principles.

However, platform accountability should not result in excessive private control over political communication. Excessive reliance on corporate moderation decisions may create another accountability problem because platforms are not democratically elected institutions. Therefore, effective governance requires balancing platform responsibility with public oversight, independent evaluation, and transparent decision-making processes. This approach recognizes platforms as important partners in democratic resilience while ensuring that political communication remains subject to democratic principles rather than purely corporate governance.

The third dimension of the framework involves institutional coordination and electoral resilience. Political deepfake governance cannot be managed by a single institution because the problem crosses multiple regulatory domains. Election management bodies are responsible for protecting electoral integrity, communication regulators oversee digital platforms, law enforcement agencies

address criminal violations, and civil society organizations contribute monitoring capacity and public education. Without effective coordination, institutional fragmentation may produce delayed responses, unclear responsibilities, and ineffective enforcement.

Collaborative governance theory provides an important foundation for understanding this challenge. Ansell and Gash (2008) argue that complex public problems require cooperation among multiple actors because no single institution possesses sufficient authority and expertise to solve them independently. In the context of political deepfakes, collaborative governance requires establishing institutional mechanisms that enable rapid information sharing, coordinated investigation, and consistent public communication during electoral periods.

For Indonesia, strengthening institutional coordination may involve creating a permanent electoral AI monitoring mechanism involving election authorities, technology regulators, academic institutions, civil society organizations, and digital platforms. Such a mechanism would allow potential deepfake incidents to be identified and assessed before they produce significant electoral consequences. Importantly, this approach would emphasize prevention and resilience rather than relying exclusively on punishment after misinformation has already spread.

The fourth dimension is democratic capacity building, which recognizes that technological solutions alone cannot eliminate the risks associated with political deepfakes. Even the most advanced detection systems cannot completely prevent manipulation because AI-generated content will continue to evolve. Therefore, democratic resilience requires strengthening citizens' capacity to critically evaluate digital information. Digital literacy and AI literacy become essential components of electoral protection because informed citizens represent the first line of defense against synthetic media manipulation.

Research on misinformation demonstrates that technological solutions are most effective when combined with public awareness,

critical thinking, and institutional trust (Lazer et al. 2018). In emerging democracies, where digital literacy levels may vary significantly across regions and social groups, public education becomes particularly important. Governments, educational institutions, media organizations, and civil society groups must therefore collaborate to develop programs that improve citizens' ability to identify manipulated content while avoiding excessive skepticism toward legitimate information.

Beyond national approaches, political deepfakes also require regional cooperation, particularly within Southeast Asia. The transnational nature of AI-generated content means that synthetic political material can be produced in one jurisdiction and distributed across multiple countries within seconds. National regulations alone are therefore insufficient to address cross-border information manipulation. ASEAN democracies face similar challenges regarding digital transformation, electoral integrity, and platform governance, creating opportunities for regional collaboration.

Regional cooperation may include information-sharing mechanisms, common standards for AI-generated political content, joint research initiatives, and cooperation between electoral institutions. While Southeast Asian countries possess different political systems and regulatory capacities, shared governance challenges create incentives for developing regional approaches. Such cooperation would strengthen collective democratic resilience while recognizing the sovereignty and institutional diversity of individual states.

The Integrated Political Deepfake Governance Framework proposed in this study therefore conceptualizes democratic resilience as the outcome of interaction among regulatory capacity, technological accountability, institutional coordination, citizen capability, and regional cooperation. This framework differs from approaches that focus primarily on technological detection or legal prohibition because it recognizes that political deepfakes represent

a systemic governance challenge rather than an isolated content problem.

The Indonesian case demonstrates why this integrated perspective is particularly relevant for emerging democracies. Indonesia possesses significant democratic resources, including electoral institutions, civil society participation, and public debate mechanisms, yet it also faces institutional limitations associated with rapid digital transformation. The challenge is therefore not simply preventing deepfake creation but developing governance arrangements capable of protecting electoral integrity while preserving democratic freedoms.

The broader theoretical contribution of this study is the argument that AI-related electoral risks should be analyzed through a governance resilience perspective rather than a technology-control perspective. Political deepfakes are likely to remain a permanent feature of digital politics because advances in generative AI will continue to expand creative and manipulative possibilities. Consequently, democratic societies must shift from attempting to eliminate synthetic media entirely toward building institutions capable of managing its risks responsibly.

In this sense, the future of democratic governance in the AI era will depend not only on technological innovation but also on institutional adaptability. Emerging democracies such as Indonesia provide valuable cases for understanding how societies can navigate the tension between technological openness and democratic protection. The Integrated Political Deepfake Governance Framework offers a conceptual foundation for future comparative research examining how different political systems develop strategies for maintaining electoral integrity in increasingly AI-mediated information environments.

V. CONCLUSION

This study demonstrates that political deepfakes represent a fundamental governance challenge for democratic systems because

they disrupt traditional assumptions underlying electoral communication, information verification, and institutional accountability. The Indonesian case illustrates that AI-generated synthetic media cannot be effectively addressed through existing misinformation regulations alone, as current legal frameworks were primarily designed to regulate conventional forms of digital manipulation rather than algorithmically generated content. While Indonesia has developed important foundations for digital governance through the Electronic Information and Transactions Law, electoral regulations, and platform governance mechanisms, significant gaps remain regarding AI-specific regulation, institutional coordination, technological capacity, and enforcement mechanisms. The findings suggest that political deepfakes should be understood not merely as a technological threat but as a broader democratic governance challenge requiring adaptive institutions capable of balancing electoral protection with freedom of expression.

This article contributes to emerging scholarship on AI governance and democratic resilience by proposing an Integrated Political Deepfake Governance Framework based on four interconnected dimensions: adaptive regulatory architecture, platform accountability, institutional coordination, and democratic capacity building. Rather than relying exclusively on legal prohibition or technological detection, the framework emphasizes ecosystem-based governance involving governments, election authorities, digital platforms, civil society organizations, technology experts, and citizens. For Indonesia and other emerging democracies in Southeast Asia, strengthening electoral integrity in the AI era requires regulatory adaptation, regional cooperation, investment in AI detection capabilities, and sustained public digital literacy initiatives. The broader implication of this study is that democratic resilience in the age of generative AI will depend less on eliminating synthetic media entirely and more on developing accountable governance systems capable of managing its risks

while preserving open political participation and democratic freedoms.

REFERENCES

Ansell, Chris, and Alison Gash. 2008. "Collaborative Governance in Theory and Practice." *Journal of Public Administration Research and Theory* 18 (4): 543–571. https://doi.org/10.1093/jopart/mum032.

Bannister, Frank, and Regina Connolly. 2014. "ICT, Public Values and Transformative Government: A Framework and Programme for Research." *Government Information Quarterly* 31 (1): 119–128. https://doi.org/10.1016/j.giq.2013.06.002.

Bennett, W. Lance, and Steven Livingston. 2018. *The Disinformation Age: Politics, Technology, and Disruptive Communication in the United States*. Cambridge: Cambridge University Press.

Chesney, Robert, and Danielle Keats Citron. 2019. "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security." *California Law Review* 107 (6): 1753–1820.

Diakopoulos, Nicholas, and Deborah Johnson. 2021. "Anticipating and Addressing the Ethical Implications of Deepfakes in the Media." *AI and Ethics* 1: 1–10.

Floridi, Luciano, Josh Cowls, Monica Beltrametti, et al. 2018. "AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations." *Minds and Machines* 28: 689–707. https://doi.org/10.1007/s11023-018-9482-5.

Gasser, Urs, and Virgilio Almeida. 2017. "A Layered Model for AI Governance." *IEEE Internet Computing* 21 (6): 58–62. https://doi.org/10.1109/MIC.2017.4180835.

Gillespie, Tarleton. 2018. *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. New Haven, CT: Yale University Press.

Gorwa, Robert. 2019. "What Is Platform Governance?" *Information, Communication & Society* 22 (6): 854–871. https://doi.org/10.1080/1369118X.2019.1573914.

Kooiman, Jan. 2003. *Governing as Governance*. London: Sage Publications.

Lazer, David M. J., Matthew A. Baum, Yochai Benkler, et al. 2018. "The Science of Fake News." *Science* 359 (6380): 1094–1096. https://doi.org/10.1126/science.aao2998.

OECD. 2019. *Recommendation of the Council on Artificial Intelligence*. Paris: Organisation for Economic Co-operation and Development.

OECD. 2024. *Facts Not Fakes: Tackling Disinformation, Strengthening Information Integrity*. Paris: Organisation for Economic Co-operation and Development.

Öhman, Carl. 2020. "Introducing the Perception Gap: A Framework for Understanding Deepfakes and Synthetic Media." *AI & Society* 35: 1–10.

UNESCO. 2023. *Guidelines for the Governance of Digital Platforms: Safeguarding Freedom of Expression and Access to Information*. Paris: United Nations Educational, Scientific and Cultural Organization.

Westerlund, Mika. 2019. "The Emergence of Deepfake Technology: A Review." *Technology Innovation Management Review* 9 (11): 39–52.