

Synthetic voices and mediated trust: a multimodal analysis of AI-generated political communication in Indonesia

Ahamed Zubair Kanakkarasan
The New College, Chennai, India

Article Info

Article history:

Received Apr 23, 2026
Revised May 21, 2026
Accepted Jun 30, 2026

Keywords:

artificial intelligence;
mediated trust;
multimodality;
political communication.

ABSTRACT

Background: Synthetic political communication increasingly detaches recognisable voices and embodied identities from biological speakers, complicating how authority, authenticity, and evidentiary credibility are established within platformed electoral environments. **Objective:** This study examines how AI-generated political communication in Indonesia constructs mediated trust through synthetic voice, multimodal orchestration, disclosure practices, and institutional contestation. **Method:** A qualitative multimodal corpus design analysed seventeen publicly verifiable documents collected between January 2023 and June 2026, comprising eight primary artefacts, four secondary verification sources, and five institutional context documents through three integrated coding instruments. **Results:** Findings show that seven primary artefacts combined synthetic speech with recognisable political likenesses, while public or direct address constituted the dominant communicative configuration. Political authority represented the most recurrent trust cue, whereas linguistic revoicing, historical revival, material promises, and apparent private access provided alternative routes to plausibility. **Implication:** Seven artefacts lacked clear AI disclosure, and every primary case subsequently attracted verification, warning, denial, or documentation from governmental, electoral, journalistic, or fact-checking institutions. **Novelty:** This study introduces an integrated framework that connects vocal identity, cross-modal congruence, platform affordances, trust cues, disclosure, communicative function, and institutional response, thereby conceptualising mediated trust as a transferable yet contested political resource rather than a measurable audience attitude within contemporary digitally mediated democratic communication systems.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Ahamed Zubair Kanakkarasan
The New College, Chennai, India
Peters Road, Royapettah, Chennai, Tamil Nadu, 600014, India
Email: ahamedzubair@thenewcollege.edu.in

1. INTRODUCTION

Indonesia's 2024 electoral cycle exposed a communication environment in which political authority could be technologically detached from speakers and redistributed through synthetic faces and voices. Mafindo documented 2,330 hoaxes during 2023, including 1,292 political hoaxes and 645 items specifically related to the 2024 election, twice the 644 political hoaxes recorded during the 2019 electoral period. Within this environment, an AI-generated video revived former president Soeharto to endorse Golkar candidates, while manipulated recordings represented President Joko Widodo speaking Mandarin and Arabic and fabricated a private conversation between Anies Baswedan and Surya Paloh. These cases are consequential not simply because they contain false or generated material, but because they appropriate recognisable vocal identities, historical memory, linguistic affiliation, and institutional status as resources for credibility.

Synthetic political communication therefore challenges verification practices centred on textual accuracy: citizens must also evaluate whether a voice, face, emotional register, and platform context constitute trustworthy political presence.

Existing research has established several components of this problem without adequately integrating them. Studies of trust in AI identify transparency, reliability, competence, and contextual expectations as central to human evaluations of automated systems [1], [2], while experimental evidence indicates that LLM-generated messages can persuade individuals on policy questions [3], [4]. Research on synthetic media maps its circulation across platforms [5] and theorises its capacity to undermine democratic deliberation, accountability, and shared evidentiary standards [6]. Work on post-human voice further shows that generative audio transforms everyday mediation and relations between speakers, technologies, and audiences [7], [8]. Meanwhile, scholarship on digital textuality, multimodal poetics, and platform-mediated meaning explains how verbal, visual, and computational modes jointly organise interpretation [9], [10], [11]. However, limited research examines how synthetic voice specifically coordinates these modes to construct mediated trust in Indonesian political communication.

Addressing this gap, this study examines a publicly verifiable corpus of AI-generated and AI-manipulated political communication circulating in Indonesia between January 2023 and June 2026. It asks three connected questions. First, how do synthetic voices reproduce political identity through vocal recognisability, linguistic revoicing, speech genre, and attributed speaker status? Second, how are these voices orchestrated with facial likeness, gesture, historical imagery, political symbols, captions, colour, layout, platform interfaces, and calls to action? Third, which dimensions of mediated trust—authority, authenticity, familiarity, benevolence, cultural proximity, or insider access—are invited, disclosed, contested, or redirected within these artefacts? To answer them, this study combines a synthetic voice and vocal identity matrix, a multimodal political orchestration grid, and a mediated trust and disclosure framework. This design distinguishes observable semiotic evidence from interpretation and institutional response, while avoiding unsupported claims about audience belief, persuasion, or behavioural effects that would require reception or experimental direct audience data.

This study advances the argument that synthetic political voices derive persuasive plausibility less from acoustic imitation alone than from cross-modal alignment among voice, embodied likeness, political memory, language choice, interface design, and circulation context. A familiar voice may index authority, but its credibility is strengthened when visual composition, captions, party symbolism, and platform conventions render technological mediation less perceptible or politically acceptable. Conversely, temporal anomalies, implausible language shifts, absent disclosure, or mismatched audiovisual cues may destabilise authenticity while increasing visibility. This proposition extends accounts of digital convergence, which emphasise that meaning emerges through interaction among media forms rather than through isolated textual channels [12], and complements research on synthetic online personae and contested authenticity [13]. Accordingly, mediated trust is treated not as a measurable audience attitude but as a semiotic and institutional accomplishment: it is invited by artefacts, negotiated through platforms, and challenged by fact-checkers, media and authorities.

2. LITERATURE REVIEW

2.1. Synthetic voice

Synthetic voice refers to speech generated or computationally altered so that vocal presence becomes separable from a biological speaker. Unlike dubbing, which openly substitutes one performer for another, synthetic voicing often preserves recognisable identity cues while obscuring mediation. Pawelec (2022) treats such media as democratic threats when deployed deceptively [6], whereas Yin et al. (2025) emphasise post-human communication in which voice becomes mediated rather than merely falsified [7]. Corsi et al. (2024) similarly show that synthetic media encompass benign, satirical, promotional, and manipulative uses [5]. Accordingly, synthetic voice is best understood as a continuum of generated presence, rather than a uniformly deceptive category.

Its analytical dimensions extend beyond technical production. Vocal form includes cloning, generation, dubbing, and hybrid manipulation; identity relation distinguishes living politicians, deceased leaders, candidates, and institutional personae; linguistic revoicing concerns shifts in language, accent, and implied audience; and authenticity cues involve timbre, rhythm, hesitation, emotional register, and apparent spontaneity. Yin et al. (2025) foreground cultural mediation [7], while Bai et al. (2025) demonstrate that generated language can acquire persuasive force independently of human authorship [3]. Studies of digital textuality also stress code, platform, and interface as constitutive elements of meaning [9], [11]. Voice analysis must integrate sound, identity, and circulation.

2.2. Multimodality

Multimodality concerns meaning produced through interaction among verbal, visual, aural, spatial, and computational modes. Foundational social-semiotic approaches view modes as culturally organised resources, whereas digital-literary scholarship increasingly treats interfaces, code, navigation, and participation as meaning-bearing structures. Assyabani (2025) emphasises visual-verbal design and screen textuality [10]; Herdiana (2025) highlights structure, agency, and interactive narrative [14]; and Mazlan et al. (2025) show that navigation itself shapes interpretation [15]. Atikurrahman (2025) further locates digital meaning within convergence across media [12]. For political synthetic media, multimodality therefore exceeds image-plus-text analysis: it includes voices, faces, captions, logos, platform conventions, and algorithmically conditioned visibility operating relationally across contemporary political contexts.

Relevant indicators include representational, compositional, interactive, and platform-specific resources. Representational features cover actors, gestures, settings, and symbols; compositional features include colour, salience, framing, layout, and logo position; interactive features involve gaze, distance, address, and calls to action; platform features include captions, hashtags, verified-looking handles, repost formats, and engagement prompts. Romadhani (2025) demonstrates how immersion reorganises emotional response [16], while Jali et al. (2025) connect transmedia design with audience engagement [17]. Hussien et al. (2025), Abd Wahid et al. (2025), and Masyithoh (2025) likewise show that digital interpretation depends on coordinated modal affordances [18], [19], [20]. Cross-modal congruence thus becomes central to perceived political authenticity judgements.

2.3. Mediated trust

Mediated trust denotes credibility constructed through technologies, signals, and performances rather than interpersonal familiarity. Afroogh et al. (2024) conceptualise trust in AI through reliability, transparency, competence, and contextual expectations [1], while Bai et al. (2025) reveal that machine-generated messages can persuade even when human authorship is absent [3]. Pawelec (2022), however, stresses erosion of democratic trust when synthetic media destabilise evidentiary certainty. Vladyslava (2025) frames authenticity as contested within synthetic influencer culture [13]. These perspectives differ over whether trust concerns confidence in systems, belief in messages, or recognition of authentic personhood. This study treats trust as discursively invited, mediated and not behaviourally confirmed.

Trust indicators can be organised into authority, competence, integrity, benevolence, familiarity, proximity, authenticity, and insider access. Disclosure adds a second dimension, distinguishing explicit labelling, implicit signalling, absent disclosure, and contested attribution. Institutional contestation adds a third, including fact-checking, denial, verification, removal, literacy intervention, and regulatory response. Research on algorithmic aesthetics and computational textuality shows that credibility is also shaped by platform form and machine-readable organisation [21], [22], [23], [24], [25]. Consequently, this study positions mediated trust as a multimodal, platformed, institutionally negotiated, and politically consequential communicative accomplishment.

3. METHOD

This study's unit of analysis comprised publicly accessible synthetic political communication connected to Indonesia, operationalised as each distinct audiovisual or audio artefact and its verified documentary record. Corpus construction produced seventeen documents: eight primary artefacts, four secondary verification or interpretive sources, and five contextual institutional documents. Primary cases included AI-generated representations of Soeharto, Joko Widodo, Prabowo Subianto, Anies Baswedan, Surya Paloh, and Rano Karno across campaign endorsement, multilingual revoicing, fabricated conversation, political allegation, and authority-based fraud. Each document received a unique identifier, while duplicate reports were linked through shared incident groups rather than counted as additional primary artefacts within this corpus.

This study employed a qualitative multimodal corpus design combining social-semiotic analysis, synthetic-voice analysis, and mediated-trust interpretation [26]. Such a design was appropriate because political deepfakes distribute meaning across sound, language, facial likeness, historical imagery, captions, interfaces, and circulation contexts rather than through isolated textual propositions. Analysis therefore treated each artefact as a structured multimodal ensemble while maintaining distinctions among observable features, interpretive categories, and institutional responses. This study did not measure audience belief or infer causal persuasion. Instead, it examined how artefacts discursively invited credibility, authenticity, authority, familiarity, benevolence, cultural proximity, or insider access through coordinated semiotic and platform-specific resources systematically.

Table 1. Composition of this dataset

Code	Actor/topic	Source category	Platform	Period	Key characteristics
PRI-001	Soeharto–Golkar	Political post documented by OECD	Instagram/web incident record	9 January 2024	Synthetic deceased president endorsing Golkar candidates
PRI-002	Joko Widodo	Komdigi verification	Social media/government notice	2023	Presidential likeness revoked in Mandarin
PRI-003	Joko Widodo	Komdigi verification	Social media/government notice	2023	Presidential likeness revoked in Arabic
PRI-004	Prabowo Subianto	Komdigi verification	Social media/government notice	2024 electoral context	Candidate represented as delivering Arabic speech
PRI-005	Anies Baswedan–Surya Paloh	Mafindo/TurnBackHoax	Social media/fact-check record	2024	Fabricated private political conversation
PRI-006	Prabowo Subianto	KPU verification	Social media/KPU article	2025	Synthetic presidential offer of financial assistance
PRI-007	Rano Karno	Komdigi verification	TikTok/government notice	Date unavailable	Fake account offering IDR20 million to followers
PRI-008	Soeharto	Komdigi verification	Social media/government notice	27 April 2025	Deceased president commenting on diploma controversy
SEC-001	Soeharto–Golkar	Metro TV	YouTube	January 2024	News discussion and reproduction of PRI-001
SEC-002	Election hoaxes	Mafindo	Website	2 February 2024	Election misinformation and AI-case overview
SEC-003	Soeharto–Golkar	OECD AI Incidents Monitor	Website	9 January 2024	Incident metadata and linked reporting
SEC-004	Soeharto deepfake	Academic journal	Journal website	2024	Prior discourse analysis of PRI-001
CTX-001	Generative AI literacy	Komdigi	Official website	January 2024	Critical-thinking and verification discourse
CTX-002	Election deepfake preparedness	InfoPublik/Kominfo	Official website	17 November 2023	Anticipatory institutional response
CTX-003	AI hoaxes and election risk	Komdigi	Official website	29 January 2024	Institutional assessment of deepfake risk
CTX-004	AI and election administration	KPU Purworejo	Official website	25 April 2025	Institutional reflection on synthetic-media threats
CTX-005	AI-based electoral supervision	Bawaslu RI	Official website	14 November 2025	Governance preparation for future elections

Information was obtained from public, legally accessible, and institutionally identifiable sources. Primary artefacts or incident-level records were traced through official political accounts, authenticated media channels, and publicly documented platform posts. Verification sources included Komdigi hoax notices, Mafindo and TurnBackHoax records, KPU publications, OECD AI Incidents Monitor entries, and national news reporting. Contextual materials comprised statements from Komdigi, KPU, Bawaslu, and InfoPublik concerning deepfakes, electoral misinformation, AI literacy, and institutional preparedness. Academic sources supported conceptual interpretation rather than case verification. Source status, publication date, political actor, platform, language, geographic scope, verification basis, and stable URL were recorded for every document systematically.

Data collection covered January 2023 to 15 June 2026 and focused on Indonesia at national and subnational levels. Searches combined Indonesian and English terms, including AI, deepfake, voice cloning, suara sintetis, pemilu, kampanye, politik, Jokowi, Prabowo, Anies, Surya Paloh, Soeharto, and Rano Karno. Items were included when AI manipulation was explicitly disclosed or verified, a political actor or institutional authority was represented, and a stable public record existed. Generic AI commentary, conventional editing, unverified reposts, nonpolitical celebrity deepfakes, and duplicate artefacts were excluded. Metadata, transcripts, excerpts, visual descriptions, timestamps, screenshots, and contextual notes were stored separately for systematic subsequent analysis.

Analysis proceeded in five stages. First, records were deduplicated by actor, message proposition, media form, approximate date, and originating incident. Second, a Synthetic Voice and Vocal Identity Matrix coded generation type, attributed speaker, language shift, speech genre, recognisability, and reported prosody. Third, a Multimodal Political Orchestration Grid examined actors, symbols, colour, framing, layout, logos, captions, platform affordances, and cross-modal congruence. Fourth, a Mediated Trust and Disclosure Framework coded authority, competence, integrity, benevolence, familiarity, proximity, authenticity, disclosure, communicative function, and institutional response.

4. RESULTS

4.1. Synthetic voice, linguistic revoicing, and political presence

Synthetic political communication in this corpus was organised principally through embodied vocal impersonation rather than disembodied audio production. Figure 1 shows that seven of eight primary artefacts combined synthetic or manipulated speech with a recognisable political face, whereas one case relied upon an apparently private audio exchange. Coding differentiated media configuration, attributed speaker status, language configuration, speech situation, and identity arrangement because each dimension contributes differently to perceived political presence. The corpus included six representations of living actors and two synthetic revivals of former president Soeharto. Five artefacts retained Indonesian, two revoiced politicians in Arabic, and one presented Joko Widodo speaking Mandarin. These distributions establish that synthetic voice operated not merely as fabricated sound, but as a structured relationship among vocal identity, political recognisability, linguistic positioning, and communicative situation.

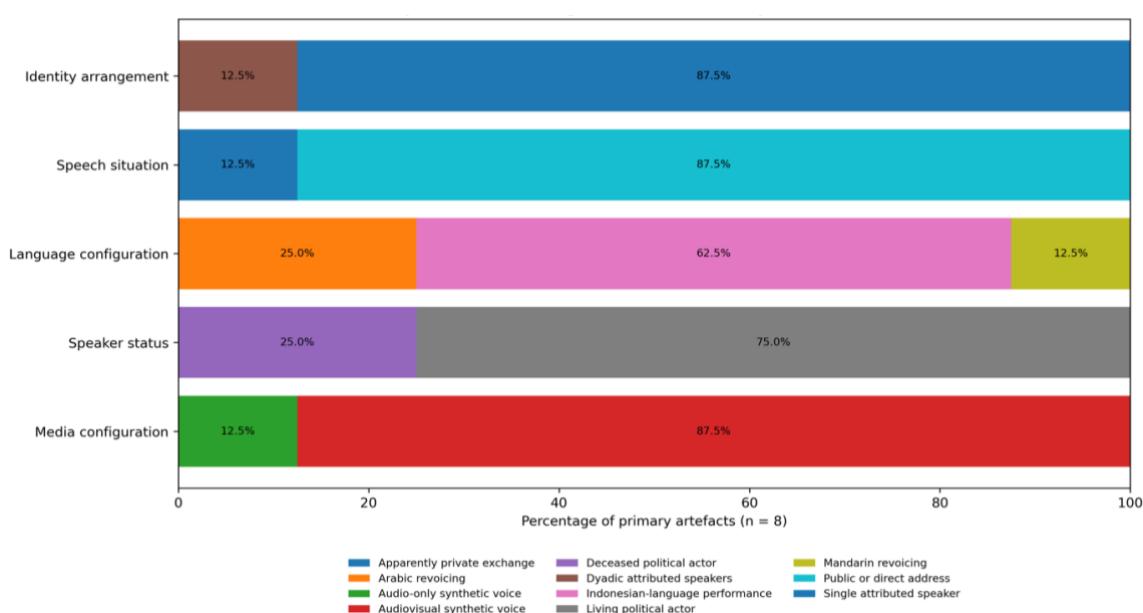


Figure 1. Synthetic-voice configurations across primary political artefacts

A dominant pattern emerged around single-speaker public address. Seven artefacts, representing 87.5% of primary data, positioned one recognisable political persona as addressing a broad or platform-mediated audience. This configuration appeared in electoral endorsement, multilingual political speech, financial-assistance solicitation, giveaway promotion, and historical testimony. Living politicians constituted 75% of represented identities, while deceased political personae accounted for 25%. Indonesian-language performance occurred in five artefacts, compared with two Arabic revoicings and one Mandarin revoicing. Only PRI-005 departed from this configuration by presenting two political actors in an apparently private conversation. Descriptively, synthetic political voice was therefore concentrated in familiar public-speaking formats, although linguistic substitution and simulated conversational privacy introduced alternative routes through which vocal credibility could be established.

These patterns support an understanding of synthetic voice as mediated political presence rather than as acoustic replication alone. Yin et al. (2025) conceptualise generative voice as a post-human form of cultural mediation in which technological production reorganises relations among identity, speech, and audience [7]. Within this corpus, recognisable political personae provided an identity anchor, while language choice and speech genre shaped how that identity became socially plausible. Arabic and Mandarin revoicing did more than alter linguistic form; they repositioned represented speakers toward imagined cultural, religious, or geopolitical audiences. Meanwhile, apparently private conversation derived authenticity from simulated spontaneity and restricted access. Consistent with Pawelec (2022), such configurations may destabilise democratic evidentiary practices because credibility is transferred from biological presence to coordinated signs of recognisability, linguistic affiliation, and communicative intimacy [6].

4.2. Cross-modal orchestration of authority and authenticity

Political authority was constructed predominantly through coordination between synthetic speech, recognisable identity, and familiar address formats. Seven of eight primary artefacts paired an attributed political voice with an embodied likeness and positioned that persona within a public or direct-address configuration. Three artefacts additionally employed marked language revoicing, while two reactivated a deceased political figure within contemporary political discourse. Platform-oriented prompts appeared in three cases, linking synthetic representation to requests for voting, contacting, following, or transferring attention toward another channel. Two artefacts connected political identity with material assistance or financial reward. Coding therefore captured representational mode, address configuration, linguistic orchestration, temporal symbolism, transactional appeal, and platform action as interconnected components rather than isolated visual or verbal variables.

Embodied direct address constituted the most recurrent configuration, appearing in 87.5% of primary artefacts. Identity–message alignment occurred at the same frequency because political personae were placed within recognisable genres such as campaign endorsement, presidential speech, testimony, solicitation, and promotional appeal. Nevertheless, five artefacts, or 62.5%, contained marked incongruities involving unexpected language, temporal displacement, or an apparently private communicative situation. Arabic and Mandarin revoicing affected three cases, while synthetic revival of Soeharto occurred in two. Material reward and assistance cues were less common, appearing in two artefacts, although both were paired with platform-specific behavioural prompts. PRI-005 remained structurally distinct because credibility depended on interactional intimacy and apparent restricted access rather than embodied public performance.

These configurations indicate that multimodal plausibility depends on both coherence and productive inconsistency. Assyabani (2025) and Nadifah (2025) conceptualise digital meaning as an interaction among visual-verbal design, code, and platform-mediated textuality [10], [11], while Atikurrahman (2025) emphasises convergence across media forms [12]. Within this corpus, recognisable faces, speech genres, historical personae, language choices, and action prompts jointly organised political meaning. Congruence strengthened attribution by making synthetic messages resemble conventional political communication; however, incongruity also generated salience by presenting politicians in unexpected linguistic or temporal roles. Historical revival transformed memory into authority, whereas welfare and giveaway messages converted political benevolence into transactional compliance. Multimodal orchestration therefore functioned not simply to imitate reality, but to redistribute authority across identity, interface, narrative genre, and platform action.

Table 2. Distribution of multimodal configurations across primary political artefacts

Analytical dimension	Category	Operational indicator	Frequency	Percentage	Cases	Data codes
Representational mode	Embodied political likeness	Recognisable face or body accompanies synthetic speech	7	87.5%	Synthetic Soeharto and Prabowo direct-address videos	PRI-001, PRI-002, PRI-003, PRI-004, PRI-006, PRI-007, PRI-008
Representational mode	Audio-led interaction	No verified embodied image; meaning is organised through attributed voices	1	12.5%	Anies–Surya Paloh fabricated recording	PRI-005
Address configuration	Public or direct address	A political persona addresses viewers or platform audiences	7	87.5%	Soeharto endorsement; Prabowo assistance message	PRI-001, PRI-002, PRI-003, PRI-004, PRI-006, PRI-007, PRI-008
Address configuration	Private interaction cue	Meaning is framed as restricted, intercepted, or backstage communication	1	12.5%	Anies–Surya Paloh conversation	PRI-005
Linguistic orchestration	Language revoicing	Visible political identity is paired with Arabic or Mandarin speech	3	37.5%	Jokowi in Mandarin; Jokowi and Prabowo in Arabic	PRI-002, PRI-003, PRI-004
Temporal-symbolic orchestration	Historical persona	A deceased political figure is reactivated in a current campaign or controversy	2	25.0%	Soeharto for Golkar; diploma controversy	PRI-001, PRI-008
Transactional orchestration	Material reward or assistance cue	Political likeness is linked to money, welfare, or material benefit	2	25.0%	Prabowo assistance offer; Rano Karno claim	PRI-006, PRI-007
Platform orchestration	Explicit platform action cue	Message requests voting, contacting, following, or payment	3	37.5%	Golkar endorsement; Prabowo assistance; Rano Karno giveaway	PRI-001, PRI-006, PRI-007
Cross-modal organisation	Identity–message alignment	Political persona corresponds with endorsement, speech, testimony, or solicitation genre	7	87.5%	Single-person synthetic political messages	PRI-001, PRI-002, PRI-003, PRI-004, PRI-006, PRI-007, PRI-008
Cross-modal organisation	Marked incongruity	Language, chronology, or interactional setting departs from expected identity context	5	62.5%	Jokowi in Mandarin; Soeharto testimony; Anies–Paloh recording	PRI-002, PRI-003, PRI-004, PRI-005, PRI-008

4.3. Mediated trust, disclosure, and institutional contestation

Trust was invited through several overlapping semiotic and institutional resources rather than through a single credibility mechanism. Authority appeared in six of eight primary artefacts, making political office, leadership, and historical presidency the most recurrent trust cue. Authenticity was coded in four cases, particularly where vocal performance, linguistic revoicing, or apparent conversational spontaneity

encouraged acceptance of represented speech. Familiarity or nostalgia occurred in three artefacts, while cultural proximity and benevolence appeared in two cases each. Insider access was concentrated in one fabricated private conversation. Disclosure was markedly uneven: only one artefact involved acknowledged synthetic production, whereas seven circulated without clear AI labelling. These distributions show that mediated trust was generally organised through recognisable political identity while technological mediation remained obscured.

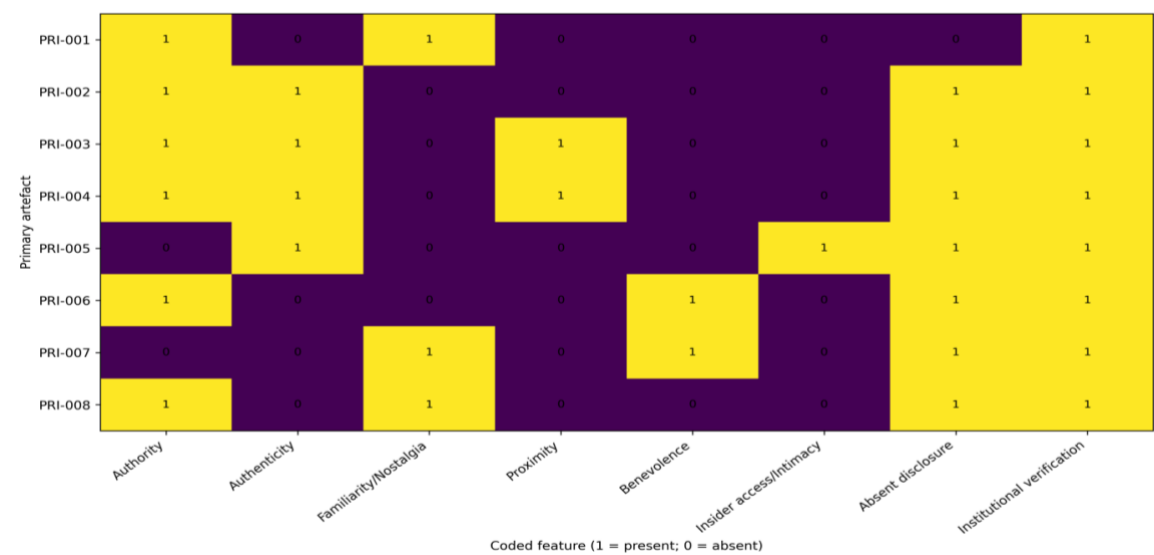


Figure 2. Trust cues, disclosure, and institutional contestation across primary artefacts

Authority dominated because six artefacts relied on presidential, candidate, party, or historical status to stabilise represented speech. Political misinformation or reputational manipulation constituted five cases, or 62.5% of primary data, whereas two artefacts used political identity for financial assistance or reward-based solicitation. Only one case operated as openly strategic campaign persuasion. Non-disclosure accompanied 87.5% of artefacts, suggesting that represented authenticity commonly depended on viewers not being immediately alerted to synthetic production. Nonetheless, all eight cases attracted institutional contestation through government verification, election-management warnings, fact-checking, incident documentation, or political denial. Descriptively, trust invitations were concentrated at artefact level, while credibility challenges emerged through secondary institutional discourse after circulation.

These findings position mediated trust as a contested communicative accomplishment rather than an inherent property of synthetic media. Afroogh et al. (2024) distinguish trust in AI through reliability, competence, transparency, and contextual expectation [1], yet this corpus shows that political artefacts frequently borrowed credibility from represented persons rather than from confidence in technology itself. Authority and familiarity transferred accumulated political recognition into synthetic messages, while benevolence and insider access supplied affective or interactional legitimacy. Absent disclosure weakened transparency but did not prevent initial plausibility, particularly when voice, face, genre, and platform cues were aligned. Consistent with Pawelec (2022), institutional verification became necessary because synthetic communication redistributed evidentiary authority across politicians, platforms, fact-checkers, and public agencies, making trust continuously negotiated rather than securely located.

5. DISCUSSION

Synthetic political voice matters because it relocates political presence from embodied speakers to reproducible signs of identity. The predominance of public-address formats indicates that synthetic communication functions by borrowing familiar genres through which authority is recognised. This can extend access, revive historical figures, or adapt messages across languages, yet it creates a dysfunction: political recognition becomes detachable from consent, presence, and accountability. Yin et al. (2025) describe generative voice as post-human mediation [7], while Pawelec (2022) emphasises its democratic risks when synthetic speech is mistaken for authentic testimony [6]. This study supports both perspectives but suggests that political danger does not arise from synthesis alone. It emerges when recognisable voice, speaker status, and communicative genre converge without adequate disclosure. Consequently, synthetic voice should be understood as an infrastructure of representational transfer: it can broaden political expression, but it can redistribute authority to actors who did not speak, consent, or participate in communication.

This pattern is intelligible because political voice operates through social expectations before acoustic precision. Audiences recognise presidents, candidates, and party leaders through accumulated associations linking timbre, speech style, public role, and communicative setting. Synthetic artefacts exploit this structure by reproducing enough identity cues to activate those associations, even when language or situation is improbable. Bai et al. (2025) show that machine-generated messages can be persuasive [3], but this corpus indicates that persuasive plausibility may depend less on machine authorship than on pre-existing symbolic capital attached to political speakers. Conversely, Corsi et al. (2024) demonstrate that synthetic media vary in function and are not uniformly deceptive [5]. This study therefore locates variation in the relationship between vocal form and institutional identity. Where synthetic speech is disclosed, recognisability may support strategic communication; where it is concealed, recognisability can simulate authorship. Underlying structure is fundamentally relational: voice gains force from political memory, status, and genre.

Multimodal orchestration has broader implications because synthetic political communication is rarely judged through sound alone. Embodied likeness, captions, language choice, historical imagery, platform prompts, and communicative genre jointly organise whether a message appears coherent, novel, suspicious, or actionable. This coordination can function productively by enabling creative campaigning and multimodal accessibility, yet it becomes dysfunctional when cross-modal coherence conceals manipulation or directs users toward fraud. Assyabani (2025), Nadifah (2025), and Atikurrahman (2025) emphasise that digital meaning emerges across visual, verbal, and computational resources rather than within a single mode [10], [11], [12]. Findings here extend that argument into political communication by showing that coherence itself can become a credibility device. Importantly, inconsistency did not simply weaken synthetic messages. Unexpected language, temporal displacement, or apparent privacy could also increase salience and circulation. Multimodal incongruity therefore performs a double role: it can expose fabrication, but it can also attract attention, invite speculation, and intensify political visibility.

These configurations arise because platformed political communication rewards recognisability, compression, and immediate interpretability. Direct address, a familiar face, a concise proposition, and a visible call to action reduce interpretive effort and make synthetic content compatible with short-form circulation. Herdiana (2025) and Mazlan et al. (2025) show that digital structures shape agency and navigation [14], [15], while Jali et al. (2025) link transmedia design to engagement [17]. This study indicates that comparable affordances operate in synthetic politics: interfaces do not merely distribute messages but organise how authority is encountered. Platform prompts translate symbolic recognition into action, including following, contacting, voting, or redirecting attention. However, this relationship should not be treated as proof of behavioural effect. Rather, platform architecture creates conditions under which political likeness, multimodal coherence, and transactional prompts can be assembled. Underlying structure is socio-technical: synthetic credibility depends on interaction among semiotic design, political familiarity, and interface logics privileging speed, salience, and engagement.

Mediated trust is consequential because this corpus shows that credibility was invited through borrowed authority rather than demonstrated reliability. Political office, historical memory, familiarity, benevolence, and apparent insider access allowed synthetic artefacts to appear meaningful before authenticity was verified. This supports Afroogh et al. (2024), who treat trust in AI as context-dependent, but it also complicates system-centred models because viewers may be responding to represented politicians rather than to AI itself. Vladyslava (2025) similarly frames authenticity as contested in synthetic persona culture [13]. The dysfunction is especially visible where benevolence or public authority is redirected into solicitation, fraud, or reputational manipulation. Yet trust cues are not inherently illegitimate; disclosed synthetic communication may use familiarity or authority transparently. What matters is how disclosure, authorship, and purpose align. Mediated trust therefore operates as a transferable resource: it can support legitimate communication, but becomes politically hazardous when recognition is detached from provenance and accountability.

This distribution can be explained by an asymmetry between rapid trust invitation and delayed verification. Synthetic artefacts mobilise instantly recognisable cues, whereas institutional responses require detection, comparison, documentation, and public correction. Pawelec (2022) argues that deepfakes threaten shared evidentiary standards [6], and findings here show how that threat is structured: authenticity is first staged through identity and genre, then contested through Komdigi, Mafindo, KPU, media, or party responses. Institutional verification therefore does more than label content false; it reconstructs provenance and reallocates authority from synthetic artefacts to trusted intermediaries. Still, such correction may not neutralise initial exposure, although this study cannot measure residual belief. Corsi et al. (2024) caution against treating all synthetic media identically, which supports differentiating disclosed campaigning, misinformation, and fraud [5]. Underlying structure is consequently temporal and institutional: synthetic trust is produced at circulation speed, while corrective trust is rebuilt through slower networks of verification, denial, literacy, and governance.

6. CONCLUSION

This study shows that synthetic political voices acquire credibility through coordinated relations among vocal recognisability, embodied likeness, language choice, historical memory, platform affordances, and institutional authority. Its principal lesson is that mediated trust is neither located solely in technological realism nor reducible to audience belief; instead, it is discursively invited through multimodal configurations and subsequently contested through verification practices. Methodologically, this study contributes an integrated framework combining synthetic voice analysis, multimodal orchestration, and trust-disclosure coding. This framework advances research by connecting acoustic identity, visual design, platform action, communicative function, and institutional response within a single comparative Indonesian political communication corpus.

This study remains limited by its small, publicly documented corpus, uneven access to original uploads, incomplete frame-level evidence, and reliance on institutional verification records for several artefacts. It therefore cannot determine audience reception, behavioural influence, longitudinal circulation, or comparative effectiveness of trust cues. Further research should expand corpus size across platforms, election periods, regions, and political actors; archive original audiovisual files systematically; and incorporate audience comments, interviews, surveys, experiments, and computational diffusion analysis. Comparative studies across languages and countries would also clarify whether authority, familiarity, disclosure, and multimodal congruence operate similarly under different political cultures, platform regimes, and regulatory environments.

ACKNOWLEDGMENTS

The author thanks the reviewers and editorial team for their constructive feedback on this manuscript.

FUNDING INFORMATION

Author states no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

Ahamed Zubair Kanakkarasan: conceptualization, methodology, formal analysis, writing – original draft, writing – review and editing.

CONFLICT OF INTEREST STATEMENT

Author states no conflict of interest.

AI USE DECLARATION

The author used artificial intelligence tools to assist with the generation of illustrative figures and with the formatting and layout of this manuscript. AI was not used to generate the substantive content, conceptualization, data, analysis, or conclusions of the research, all of which remain the sole responsibility of the author.

INFORMED CONSENT

Not applicable – this study did not involve direct human participants. Data consisted solely of publicly verifiable synthetic-media artefacts and institutional documentation.

ETHICAL APPROVAL

This research did not involve human or animal subjects; it analyzed publicly accessible synthetic-media artefacts and institutional documentation, and therefore no institutional review board approval was required. Where research does involve human subjects, the author affirms that such research would be conducted in full compliance with all relevant national regulations and institutional policies, in accordance with the tenets of the Declaration of Helsinki, and would be approved by the author's institutional review board or equivalent committee.

DATA AVAILABILITY

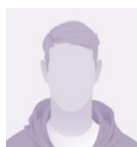
The corpus of documents analyzed in this study are publicly accessible and available from the corresponding author upon reasonable request.

REFERENCES

- [1] S. Afroogh, A. Akbari, E. Malone, M. Kargar, and H. Alambeigi, "Trust in AI: progress, challenges, and future directions," *Humanities and Social Sciences Communications*, vol. 11, Mar. 2024, doi: 10.1057/s41599-024-04044-8.
- [2] A. Fawaid, P. Handayani, and Y. A. Abdillah, "E-Portfolio in Improving Critical Thinking and Self-Management through Lesson Study: A Study on Writing Pedagogy in Higher Education," presented at the 2024 10th International Conference on Education and Technology (ICET), IEEE, 2024, pp. 149–154. doi: 10.1109/ICET64717.2024.10778453.

- [3] H. Bai, J. Voelkel, S. Muldowney, J. Eichstaedt, and R. Willer, "LLM-generated messages can persuade humans on policy issues," *Nature Communications*, vol. 16, Jul. 2025, doi: 10.1038/s41467-025-61345-5.
- [4] A. Fawaid, M. N. F. Hidayat, A. Afifudin, R. A. Setiawati, L. H. Awaliyah, and H. S. Malika, "Integrating Decision Aid Science into Game-Based Interactive Textbooks: Improving Students' Engagement and Conceptual Understanding in Indonesian Language Classes," *KEMBARA: Jurnal Keilmuan Bahasa, Sastra, dan Pengajarannya*, vol. 12, no. 1, 2026, doi: 10.22219/kembara.v12i1.42972.
- [5] G. Corsi, B. Marino, and W. Wong, "The spread of synthetic media on X," *Harvard Kennedy School Misinformation Review*, Jun. 2024, doi: 10.37016/mr-2020-140.
- [6] M. Pawelec, "Deepfakes and Democracy (Theory): How Synthetic Audio-Visual Media for Disinformation and Hate Speech Threaten Core Democratic Functions," *Digital Society*, vol. 1, Aug. 2022, doi: 10.1007/s44206-022-00010-6.
- [7] S. Yin, P. Xie, and Y. Niu, "The Post-Human Voice in China: Generative AI, Cultural Mediation, And the Transformation of Everyday Communication," *International Journal on Artificial Intelligence Tools*, Nov. 2025, doi: 10.1142/s0218213025500198.
- [8] A. Fawaid, R. Assyabani, I. Abdullah, C. Muali, M. S. Itqan, and S. Islam, "Human Intelligence and Algorithmic Precision: An Experimental Study of Indonesian Translation Pedagogy in Higher Education," *Asian Journal of University Education*, vol. 21, no. 3, pp. 779–792, 2025, doi: 10.24191/ajue.v21i3.53.
- [9] A. Fawaid, "Mapping the field of digital literary studies: Concepts, methods, and emerging directions," *Lingua Technica: Journal of Digital Literary Studies*, vol. 1, no. 1, pp. 1–13, Jan. 2025, doi: 10.64595/99c6t274.
- [10] R. Assyabani, "Multimodal poetics in digital literature: A corpus-based analysis of visual-verbal design, screen-based textuality, and reader interaction," *Lingua Technica: Journal of Digital Literary Studies*, vol. 1, no. 1, pp. 41–53, Jan. 2025, doi: 10.64595/cy0q2n14.
- [11] S. A. Nadifah, "Language, code, and platform-mediated textuality in electronic literature," *Lingua Technica: Journal of Digital Literary Studies*, vol. 1, no. 1, pp. 28–40, Jan. 2025, doi: 10.64595/80xngp06.
- [12] M. Atikurrahman, "Reimagining textuality: Digital convergence and literary adaptation in Indonesia," *lingtech*, vol. 1, no. 2, pp. 63–71, Jun. 2025, doi: 10.64595/lingtech.v1i1.30.
- [13] S. Vladyslava, "AI and Deepface Influencers: The Challenge of Authenticity in the Online Space," *Universal Library of Engineering Technology*, Jun. 2025, doi: 10.70315/uloap.ulete.2025.0202004.
- [14] N. Herdiana, "From hypertext to interactive narrative: Structure, agency, and multimodal meaning-making of e-literature," *Lingua Technica: Journal of Digital Literary Studies*, vol. 1, no. 1, pp. 14–27, Jan. 2025, doi: 10.64595/gazjdj89.
- [15] N. H. Mazlan, C. W. Putra, and H. Sulisty, "Understanding reader navigation patterns in multi-path hypertext fiction: A case study approach to Patchwork Girl," *lingtech*, vol. 1, no. 2, pp. 1–12, Jun. 2025, doi: 10.64595/lingtech.v1i1.3.
- [16] A. D. Romadhani, "Virtual reality as narrative medium: The emotional effects of full immersion in VR-based film Aladin," *lingtech*, vol. 1, no. 2, pp. 51–62, Jun. 2025, doi: 10.64595/lingtech.v1i1.29.
- [17] H. bin Jali, M. Mawaidi, and L. Habiburrahman, "Transmedia storytelling in expanding audience's engagement: A phenomenological study of Nussa Rara," *lingtech*, vol. 1, no. 2, pp. 25–37, Jun. 2025, doi: 10.64595/lingtech.v1i1.20.
- [18] B. T. A.-Z. Hussien, C. I. Liyana, and N. Nadifah, "Digital poetics in the classroom: The role of Taroko Gorge in increasing students' engagement and critical thinking skills," *lingtech*, vol. 1, no. 2, pp. 13–24, Jun. 2025, doi: 10.64595/lingtech.v1i1.4.
- [19] N. A. Wahid, F. Farhan, and M. Wijaya, "The role of interactive gamification through Wordwall Apps to improve students' engagement and literary comprehension," *lingtech*, vol. 1, no. 2, pp. 38–50, Jun. 2025, doi: 10.64595/lingtech.v1i1.21.
- [20] D. Masyithoh, "Teaching digital literature in secondary school: Multimodal meaning-making and students' interpretive engagement," *Lingua Technica: Journal of Digital Literary Studies*, vol. 1, no. 1, pp. 67–79, Jan. 2025, doi: 10.64595/5k8np621.
- [21] Y. A. Abdillah, "Poetics of algorithmic excess: digital aesthetics in Indonesia's Twitter poetry bot," *Lingua Technica: Journal of Digital Literary Studies*, vol. 2, no. 1, pp. 86–101, 2026, doi: 10.64595/lingtech.v2i1.138.
- [22] K. A. Anindita, S. Hockey, and T. W. Septianto, "Mapping thematic patterns in Indonesian novels through concept mining and computational linguistics," *Lingua Technica: Journal of Digital Literary Studies*, vol. 2, no. 1, pp. 1–17, 2026, doi: 10.64595/lingtech.v2i1.128.
- [23] R. W. Ikomah and Z. H. Sain, "Text mining and semantic modeling of literary corpora: a machine learning-based study of Indonesian fiction," *Lingua Technica: Journal of Digital Literary Studies*, vol. 2, no. 1, pp. 51–67, 2026, doi: 10.64595/lingtech.v2i1.133.
- [24] M. N. Kumar and M. A. Yaqin, "From manuscript to metadata: preserving Indonesian literary heritage in the contemporary digital transition," *Lingua Technica: Journal of Digital Literary Studies*, vol. 2, no. 1, pp. 18–32, 2026, doi: 10.64595/lingtech.v2i1.130.
- [25] M. R. M. Nor and A. Zubaidi, "Digital reading platforms in literary learning: cultural understanding and engagement in Indonesian and Malay texts," *Lingua Technica: Journal of Digital Literary Studies*, vol. 2, no. 1, pp. 33–50, 2026, doi: 10.64595/lingtech.v2i1.131.
- [26] J. W. Creswell, *Research Design, Qualitative, Quantitative and Mixed Method Approach, (Research Design, Pendekatan Kualitatif, Kuantitatif, dan Mixed, Edisi Ketiga) terj. Achmad Fawaid*. Yogyakarta: Pustaka Pelajar, 2011. [Online]. Available: <https://digilib.ab.ac.id/opac/detail-opac?id=139455>

BIOGRAPHIES OF AUTHORS



Ahamed Zubair Kanakkarasan     is an Associate Professor at The New College, Chennai, India. His academic background is in language teaching and pedagogy, with scholarly interests extending to multimodal communication and digital/political discourse. He can be contacted at email: ahamedzubair@thenewcollege.edu.in