



DOI: [https://doi.org/10.71317/kjard.2.8\(s\).2026.530](https://doi.org/10.71317/kjard.2.8(s).2026.530)

The Kashmir Journal of Academic Research and Development

Journal homepage: <https://rjsaonline.org/index.php/KJARD>



Deepfakes, AI-Driven Disinformation, and Global Politics: Emerging Challenges for Diplomacy, Elections, and International Security

Muhammad Irfan Magray

IR Department, Alhamd Islamic University, Quetta Campus

Email: magrayo6@gmail.com

Shakeel Shaheen

IR Department, Alhamd Islamic University, Quetta Campus

Email: termuthain@gmail.com

Dr. Nadia Shaheen

Assistant Professor, Alhamd Islamic university Islamabad campus

Email: snadia@ir.qau.edu.pk

ARTICLE INFO

ABSTRACT

Received:

July 09, 2026

Revised:

July 22, 2026

Accepted:

August 08, 2026

Available Online:

August 21, 2026

Keywords:

Deepfakes, synthetic media, disinformation, computational propaganda, electoral integrity, diplomacy, international security, political legitimacy, information integrity, AI governance

Corresponding Author:

Muhammad Irfan Magray

Email: magrayo6@gmail.com

Generative artificial intelligence (AI) and synthetic media are transforming the modern information landscape by facilitating the creation and dissemination of text, audio, images and video that are highly realistic and can be created and reproduced at a rapid pace. Deepfakes" – sound and video recordings, or multimodal content, created or manipulated using AI to represent real people, events, or authentic statements – are among these developments and pose specific challenges to international politics, as they can create the illusion of real evidence and actual representation. Misinformation can be false or misleading information disseminated without necessarily intending to mislead, whereas disinformation involves deliberate attempts at misleading, malinformation uses real information used in a harmful way and deepfakes involve the use of synthetic media which could be part of a wider computational propaganda and AI-driven political manipulation. The focus of this work is on the impact of these technologies on the nature of diplomacy, elections and international security, as well as on political legitimacy, public trust and the stability of information ecosystems. Using an interpretivist qualitative research approach, the study uses qualitative document analysis and thematic analysis of academic literature, policy documents, documented political cases, institutional reports and other authoritative documents. Themes emerging from the analysis include interrelated issues related to the breakdown of epistemic trust, the increasing speed and personalisation of political influence, the use of uncertainty strategically and the challenge of verifying disputed digital content. Synthetic depictions of political leaders can add to official communication misunderstanding, exacerbate misperceptions, make it harder to signal crises and create more space for foreign interference. In electoral contexts, AI can be used to enhance impersonation, voter manipulation and reputational attacks, and to create coordinated disinformation and undermine confidence in electoral institutions and outcomes (International IDEA, 2024). Deepfakes can be used in international security as tools for psychological operations, information warfare, and strategic deception, especially in times of geopolitical tension. The impact of their work is unpredictable and not pre-determined, however; it is influenced by political polarisation, views of the audience, institutional capacity, the media environment, the platform architecture, and the acceptance of information sources. The study adds to the fields of international relations and political communication scholarship because conceptually, the study does not view deepfakes as simply a technology but as tools that can shift the environment in which political actors process evidence, build credibility and act collectively. These results suggest that the detection tools are not enough to achieve effective responses; they need international collaboration, institutional resilience, media and AI literacy, transparent governance of platforms, verification and provenance tools, and responsible governance of AI to safeguard information integrity without stifling legitimate political expression. The study thus calls for a multi-dimensional governance structure that seeks to tackle the technological production of synthetic content and the political context in which it is produced. The research focuses on what is known as the "liar's dividend," or how synthetic media can help recordings to be dismissed as fake, thereby reducing accountability and leading to ambiguity (Chesney & Citron, 2019). This refocuses attention from detecting falsehoods to supporting processes for establishing authenticity. It places standards upon the states, authorities, journalists, platforms, civil society, and institutions in relation to maintaining expression. Information integrity must be considered a cross-national issue.

Introduction

Generative artificial intelligence (AI) is reshaping the modern global information landscape. Technological and economic access barriers to persuasive digital content have significantly decreased due to advances in machine learning, automated content generation, synthetic media and platform-based communication. The landscape of political communication is thus changing in a context of images, voices, videos, and statements that can be created and manipulated at unprecedented speed and scale. This is an opportunity for communication and participation but also challenges some assumptions regarding authenticity, evidence, and credibility. This increasing visibility of generative AI, in turn, is a new political phenomenon and not only a technological evolution that will have impact on democratic governance, international relations, and security (Bennett & Livingston, 2018; UNESCO, 2023).

In this evolving landscape, the deepfake is a very prominent synthetic media type. Deepfakes are AI tech tools that can create or alter AV material, allowing people to sound or look like they are saying or doing things that they never actually did. The political relevance of their work comes from their ability to mimic familiar faces and familiar events, and to make use of the trustworthiness traditionally accorded to visual and auditory evidence. It has been proven that fake media can make it more difficult for viewers to tell real media from fake media (Vaccari & Chadwick, 2020) and that their persuasive impacts differ depending on the media context, political views, characteristics of the media message, and the sensitivity of the audience. The central political question, in other words, is not only if a specific deepfake can fool an audience, but also if the growing ease of creating believable synthetic content changes the political environment in which political actors gain credibility and build trust.

Deepfakes also differ from other types of information disorder. While misinformation is the spread of false or misleading information without necessarily being done with the intent to deceive, disinformation is the strategic production or dissemination of false or misleading information. Malinformation is the use of authentic information for harmful purposes, and synthetic media is more generically the creation of and manipulation with authentic information. Deepfakes can be used in the context of disinformation and computational propaganda, influence operations, or political manipulation, but it is important not to conflate these terms with the Deepfakes. The concept difference is crucial because the potential for political impact is not just a product of the technical nature of synthetic content, but also of the institutional settings and the networks, audiences and intentions in which it moves.

Deepfakes are significant not just on a domestic scale but on an international scale as well. In the context of elections, political videos that are manipulated can enable voter manipulation, reputational attacks, impersonation and coordinated influence efforts, thereby impacting the confidence in candidates and electoral institutions. However, it would be analytically challenging to take the view that all deep fakes effect voting behavior or electoral outcomes. Previous studies indicate that it is not automatic; instead, audiences' interpretations of manipulated information are affected by political polarization, pre-existing beliefs, the credibility of the source and the type of media environment (Vaccari & Chadwick, 2020). What may be a more basic question, however, is the steady loss of faith in the informational basis of democratic decision making. The doubts of citizens about the authenticity and legitimacy of political evidence can increasingly make political disagreement and dispute over the actual fact of the events happen at the same time.

This is a very serious issue for diplomacy and international security. The ability to send credible signals, verify statements and interpret the politics in time are crucial to the diplomatic process. A phony recording of a political leader, military commander or government body might add to the confusion in a sensitive crisis, making diplomatic signaling more challenging and introducing more opportunities for strategic deception. State and non-state actors could also use deepfakes to shape perceptions beyond borders as a tool in information warfare and foreign influence operations. But the potential for security of such technologies ought to be seen as a strategic possibility and institutional vulnerability, not as an inevitability when synthetic media is created. The effects will probably vary based on verification capacities, political ties, media infrastructure, and the stability of institutions where crisis communication and public information is handled.

One of the relevant aspects of this challenge is epistemic uncertainty. Deepfakes are not only a way to add false representations to the information landscape, but can also erode trust in representations. This is linked to the notion of the "liar's dividend," in which those who are being accused of misconduct can call into question any truthful evidence because audiences are becoming aware that they can be fabricated (Chesney & Citron, 2019). This creates an asymmetric problem which is: society has to cater for the case where they find themselves believing a fabricated fact, but also the case where they find themselves disbelieving a true fact. This uncertainty can have a positive effect on those political players who want to get away with it, and can undermine the evidentiary basis of journalism, diplomacy, democratic controls, and international crisis management.

While there has been significant scholarship on misinformation, computational propaganda, synthetic media, and deepfakes, there is still a critical conceptual gap. While there are some valuable insights to be taken from existing research, these aspects are often studied in isolation and do not always connect explicitly to electoral manipulation or the public's reactions to misleading information. The strands of existing research have yielded valuable insights into electoral manipulation, public

reactions to deceptive media, platform governance, foreign influence, and information warfare but these have not always been connected directly to electoral manipulation or to public reactions to misleading information. However, the study of deepfakes' relationship with domestic political communication and transnational influence operations and international security dynamics has been neglected. This research study aims to fill that void and explores deepfakes as a multi-faceted political phenomenon and its implications in terms of technology, institutions, political actors, audiences, and information environments.

This is a qualitative study that asks the question: In what ways are deepfakes changing the face of political communication and power dynamics in modern international politics? What impact do deepfakes have on citizens' trust in politicians, electoral integrity and democratic legitimacy? How might be the challenges of diplomatic communication, crisis signaling and international security be complicated by synthetic media? What are the ways in which state and non-state actors have or might be using/deploying deepfakes in the context of broader disinformation and influence operations? Lastly, how might forms of institutional and international governance be used as a mechanism to build resilience to the political manipulation of AI, while simultaneously safeguarding political expression?

The study uses a qualitative interpretivist paradigm and qualitative document analysis and thematic analysis methods are used to analyze the relevant academic literature, policy documents, institutional documents, and documented cases of politics. It conveys the importance of tying together discussions often siloed in IR, political communication, democratic governance, election security, diplomacy studies and AI governance. The study is not only about the technical problem of the detection of deepfakes, but also about the evolving political information landscape, in which credibility, authenticity and political communication are increasingly called into question. The article thus is a contribution to a more holistic understanding of the potential of synthetic media on political legitimacy, diplomacy, and international security. The rest of the article examines existing literature, clarifies the qualitative approach, summarizes key findings and themes, explores theoretical and policy implications, and explores avenues for institutional resiliency, responsible AI governance, and enhanced international cooperation.

Literature Review

The field of deepfakes has grown organically across the fields of artificial intelligence, political communication, information disorder, and security studies. Deepfakes can be best defined as an AI-generated or AI manipulated audiovisual representation of persons, events, or statements that did not actually take place. They are a subset of synthetic media, and generative AI is the underlying technology that allows this type of content to be generated at scale. Deepfakes do not have the same meaning as misinformation or disinformation, at least conceptually. Misinformation is any information that is factually incorrect, even if it is not meant to mislead, while disinformation is information that is deliberately false and deceptive, and malinformation is information that is harmful, but accurate. Synthetic content could be used by political or state actors in computational propaganda and information operations. The political importance of a deep fake is more than the fact that it is a fake; it depends on how actors create, share, interpret, and use deepfake videos in a way that may be beneficial to their goals.

As the technology evolves, more and more research is pointing to how deepfakes are used in tandem with uncertainty, not as a persuasive tool. For instance, Vaccari and Chadwick (2020) were able to show that political deepfakes might have the power to create uncertainty more easily than the ability to deceive, and that such uncertainty would lead to less trust in social media news. This puts to the test technologically deterministic perspectives that imply, if not insist, that audiences will inevitably be convinced by realistic synthetic video. Even more recently, it has been found that the impact of manipulated political content is contingent upon plausibility, ideological congruence, source credibility, and contextual situations. In a recent study in the *Journal of Memory and Behavior*, Hameleers et al. (2024) show that plausible manipulations can seem more believable, and even if they are not, deepfakes of political figures can still seem to delegitimize them. Therefore, deepfakes need to be seen in the context of an emerging multipolitical communication field in which the authenticity, credibility, and interpretation of images is increasingly challenged.

This is reflected in the electoral literature. Deepfakes can be used to impersonate candidates, create fake speeches, deceptive ads, reputation attacks, and targeting voters. It has been shown in documented instances that AI-generated political audios and videos have already infiltrated into electoral communication, especially via social media platforms and quickly circulated campaign content (Łabuz & Nehring, 2024). However, there is no evidence that the use of deepfakes will invariably change election results. Their impact can be indirect via confusion, disruption of the agenda, polarisation, loss of institutional confidence and reinforcement of existing political narratives. In this regard, the "liar's dividend" is especially significant: Once they know that there are convincing lies, they can declare the real evidence to be fake (Chesney & Citron, 2019). Therefore, deepfakes have the potential to undermine democratic accountability even if viewers know that the information is fake. As a result, deepfakes can erode democratic accountability whether or not they are viewed as true.

The implications for diplomacy and international communication are discussed in a related literature. Credible signaling, interpretation and attribution, coupled with trust in official communication, are key to diplomatic relations. Synthetic audio or

video of political leaders' statements, intentions or actions could then raise confusion. Even in crisis situations, uncertainty can make it hard to show up where it's needed and make it harder to tell which signals are deliberate and which are manipulative. The international context is particularly relevant as AI-generated media products can quickly move across borders, blurring traditional boundaries between domestic political communication and foreign information operations. Deepfakes can then be weapons that states and non-state actors use to shape international perception, discredit rivals or try to sway the public opinion of political events.

The concept of synthetic media in security scholarship is also presented as a means of information warfare, psychological operations, hybrid warfare, and strategic deception. The worry cannot only be around fabricated political statements: distorted depictions of military operations, government announcements or crisis events may have a significant impact on decision making in times of tension. However, there are warnings in the literature that all synthetic artifacts are not directly a threat to security. The impact of misinformation will be strategic, depending on the timing, who the audience is, how much institutional trust is given to those disseminating the information, the political climate and the ability of authorities to check the information. However, deepfakes are not an independent type of warfare, but are rather a new dimension in information operations.

The developing idea of epistemic security can help to connect political communication to international security. When citizens, journalists, diplomats and policy makers can't be sure that digital evidence is accurate, the situation is larger than just one's being deceived. It is about the trustworthiness of information in the environment. This concern is backed up by recent research on human detection: some manipulated content can be detected, but the performance is dependent on the modality and the technological sophistication (Barari et al., 2024). Therefore, the technology-based solution is not a complete answer.

Governance scholarship is becoming more and more proponents of layered systems, which include detection, provenance, watermarking, fact-checking, media literacy, platform accountability, and institutional verification. There are a number of successful approaches to provenance, including Content Credentials, that do convey information about the creation or modification of digital information, and there are systems to detect manipulation after the fact. Both of these methods are not enough on their own, however. Metadata can be stripped away, alarm systems can be out of date, and labelling can even be confused. Therefore, the role of effective governance demands coordination between platforms, between governments, between journalists, between technology companies, between civil society and between international institutions.

In general, the literature shows a high degree of disunity. The previous scholarship has produced a lot of knowledge in the areas of electoral manipulation, political persuasion, misinformation, detection, platform governance and information warfare, but these are studied individually. What is not adequately understood is the concept of deepfakes as a multi-lingual and multi-jurisdictional political phenomenon which connects electoral legitimacy, diplomatic communication, epistemic uncertainty and international security. This study fills that gap by conducting a qualitative interpretivist synthesis of these interrelated fields; deepfakes are not just technologies, but tools that can transform the way political actors create credibility, communicate strategically, and challenge political reality.

Research Methodology

The study is based on a qualitative and interpretivist approach to address the implications of deepfakes and AI generated disinformation for diplomacy, elections, political communication and international security. Given the study's focus on meaning-making around authenticity, credibility, deception and security rather than frequency or statistical effects, the paradigm used is that of interpretivism. Interpretivist research acknowledges that political realities are not in itself, but rather are created by narratives, interpretation, and the social contexts in which they are produced (Creswell & Poth, 2018). This is especially important when it comes to deepfakes studies as the meaning of the deepfake content relies on the audience's interpretation of the synthetic content and the way institutions deal with uncertainty.

A qualitative exploratory research design is used. Exploratory qualitative inquiry would be suitable for a phenomenon that is new or growing, and has not reached a stage of fullness of understanding (Merriam & Tisdell, 2016). Deepfakes are a new and evolving technology and political phenomenon that has implications for communication, elections, diplomacy, cybersecurity, and strategic competition. A qualitative design then allows the study to explore the relationships among these dimensions and the identification of patterns that would not be perceived by an approach that would only consider measurable effects.

The study is based on documentary and textual data rather than interviews, questionnaires, experiments and statistical data. The corpus includes peer-reviewed journal articles, academic books and chapters, policy reports, government documents, documents from international organisations, election monitoring reports, diplomatic and foreign-policy documents, cybersecurity and security studies reports, credible institutional publications and documented cases on political deepfakes and AI-driven disinformation. The sources cited are directly relevant to the research questions, are scholarly or institutionally credible, offer substantive empirical or conceptual material, and are relevant to one or more dimensions of deepfakes, political communication, elections, diplomacy, and/or security. Opinion pieces that do not have evidentiary foundations, that have been

duplicated from elsewhere, that are unverifiable from social media and that have no meaningful link to the research problem are not included.

Sources are selected with a purposive sampling strategy. Instead of sampling for number, purposive sampling allows the researcher to find materials that have information-rich characteristics that can be used to illuminate the phenomenon under study (Patton, 2015). Sources are chosen relevant to the topic, authoritative, analysis, method and direct relevance to the research questions. Care is taken to include diversity across academic disciplines and institutional perspectives in order to avoid a technological or political interpretation only.

Thematic analysis is the main analytical approach used. Building on Braun and Clarke's (2006) flexible, systematic approach, the analysis is carried out in a few interrelated steps. The researcher first gets to know the document(s) by repeating readings and studying them in context. Second, passages that are relevant to the study are given codes to help identify important concepts, arguments, cases, and institutional responses. Third, ideas are related back to other ideas within the texts to find general trends. Fourth, codes that are related are grouped into candidate themes. Fifth, themes are examined and developed for internal consistency and meaningful differences between the categories. Sixth, the interrelationship of themes is explained in terms of research questions and the broader theoretical literature. Lastly, the themes are viewed as a whole and developed as an integrated qualitative interpretation instead of as separate findings.

The analysis is particularly focused on epistemic uncertainty, electoral manipulation, diplomatic credibility, foreign interference and strategic deception, political polarization, institutional trust, international security, governance, and democratic resilience. Themes serve as an analytical connection between the domestic and international aspects of AI-driven disinformation. The study thus examines not only the potential of deepfakes to mislead but also the impact that their existence could have on the way evidence, credibility, attribution, and political legitimacy is perceived.

Trustworthiness in qualitative studies is achieved by three steps of credibility, dependability, and confirmability (Lincoln & Guba, 1985), and transferability. Source triangulation reinforces credibility by making comparisons of claims from academic, governmental, institutional, and documented case materials. This is achieved by clear documentation of principles and stages of analysis for the selection of sources. The confirmability is increased when interpretations are based on identifiable document evidence and not on researcher assumptions. The transferability of the results is facilitated by extensive descriptions of contexts, themes and analytical processes, enabling readers to make judgements about the relevance of the results to similar political contexts (Tracy, 2010).

Ethical issues are especially salient due to the use of potentially harmful synthetic media and politically charged misinformation in the subject matter. The study therefore emphasises the proper representation of evidence, proper and responsible citation, contextualisation of documented evidence, and the avoidance of unnecessary reproduction and amplification of harmful material fabricated. Just because a claim is floating around the Internet does not make it true. The analysis is also free from technologically deterministic or politically sensational reading.

Last, the study has some limitations. Some of the findings are time-sensitive, as the situation with deepfakes changes rapidly; it is difficult to make the case for the existence of documented cases of misuses of deepfakes in a uniform manner across countries and political systems; algorithms and the moderation of platforms are often opaque; and documentary research is unable to establish unambiguous causal links between individual cases of deepfakes and political outcomes. Geographical and Western institutional bias may also exist in the current scholarship. As such, the study does not assert that deepfakes necessarily lead to electoral disruption, political instability or international conflict. Rather, it catalogues, sharpens, and documents interpretive patterns, institutional weaknesses, and potential pathways for the impact of synthetic media on current politics on the planet.



Findings and Discussion

The five themes from this thematic analysis of academic literature, policy documents, institutional reports and documented political cases were the transformation of political truth, electoral and democratic vulnerability, erosion of evidentiary trust, diplomatic and international-security risks, and governance and democratic resilience. Instead of claiming that deepfakes are inevitably convincing or decisive in political outcomes, the results suggest that the broader implications of deepfakes are the way in which political information is created, understood, verified, and challenged.

Theme1: examines how the concept of 'political truth' is changing with deepfakes

The first theme is manifested in the destabilization of audiovisual authenticity. The documentary evidence suggests that the conventional view of the reliability of what one sees and hears of a political actor is being undermined by the deepfakes. The conclusion that audiences automatically accept synthetic material as "real" however, is not supported by the evidence. Results indicate that political deepfakes have effects that vary based on context, such as plausibility, political dispositions, credibility of the source, and the nature of the manipulated information (Vaccari & Chadwick, 2020). Likewise, a study of human detection shows that there is some capacity to recognize manipulated material, but this is not perfect, and it depends on the type of content (Nightingale & Farid, 2022).

The key qualitative result then is that the political relevance of deepfakes cannot be explained just as a deception tool. Their more consequential impact is likely to be the generation of epistemic uncertainty: uncertainty about the trustworthiness of political evidence, whether political evidence can be trusted at all. This is an expansion on the existing disinformation studies paradigm, which has been focussed on users' perception of false information, to a focus on the maintenance of trust in mechanisms for truth.

Table 1. Deepfakes and the Transformation of Political Truth

Qualitative finding	Supporting pattern in the literature	Political interpretation	Implication
Deepfakes weaken assumptions about audiovisual authenticity	Increasingly realistic synthetic audio and video challenge conventional evidence assessment	Visual evidence becomes contestable	Verification becomes increasingly important

Deepfakes do not automatically persuade audiences	Effects vary according to context, plausibility, source and audience	Technological determinism is insufficient	Political effects should be assessed contextually
Synthetic content can generate uncertainty	Political deepfakes may increase uncertainty even without direct belief	Epistemic uncertainty becomes politically consequential	Trust-building institutions become essential
Authentic and fabricated material may become difficult to distinguish	Detection performance is imperfect	The boundary between evidence and fabrication becomes contested	Provenance and authentication mechanisms gain importance

Note. Synthesized from the qualitative analysis of relevant literature, including Vaccari and Chadwick (2020) and Nightingale and Farid (2022).

Theme 2 Electoral Integrity and Democratic Legitimacy

The second theme is about the fragility of electoral communication. In the literature, the themes of concern are candidate impersonation, fabricated statements, manipulated campaign materials, reputational attacks, and misleading political advertisements. Some of the technical challenges which have been linked to advanced AV manipulation are being mitigated by the increasing accessibility of generative AI. The existing cases of political content manipulation show that artificial content has already been part of electoral communication in various countries of the world (Łabuz & Nehring, 2024).

However, a big caveat is necessary to the findings. At this point, there is no evidence showing that DFFs necessarily influence voters' decisions or election outcomes. Rather, their political impacts could be indirect via confusion, reputational damage, polarization, disruption of the political agenda, and loss of confidence in candidates or electoral institutions. It is significant because democratic legitimacy is not only a matter of the actual integrity of electoral processes, but also of perceptions of them and of their results as credible.

Table 2. Electoral Integrity and Democratic Legitimacy

Dimension	Qualitative finding	Potential political effect	Evidence assessment
Candidate impersonation	Synthetic content can imitate political figures	Reputational damage and voter confusion	Demonstrated capability
Fabricated statements	AI can produce apparently authentic political speech	Campaign narratives may be distorted	Demonstrated technological risk
Political polarization	Manipulated content can reinforce existing partisan narratives	Greater distrust between political groups	Context-dependent
Electoral legitimacy	False content may generate disputes about candidates or results	Reduced confidence in democratic institutions	Plausible and context-dependent
Voting behaviour	Direct influence is difficult to establish	Potential persuasion or demobilization	Evidence remains limited

Note. The table distinguishes documented technological and political uses from effects that remain context-dependent rather than assuming that deepfakes automatically change electoral outcomes.

The results thus indicate that the information integrity aspect of the elections should not be neglected in election security strategies, along with conventional aspects of cyber security. Electoral authorities need to have rapid means of verifying official communications and political parties, journalists, and civil society groups need reliable channels of verifying manipulated info.

Theme 3: The Liar's Dividend and the Attack on evidentiary trust

The third theme is related to the "liar's dividend." Genuine evidence is available and with convincing synthetic media, it is possible to contest the evidence as being manipulated. The advent of deepfakes may thus create a secondary issue, along with the possibility of direct deception, of making "real" evidence easier to disbelieve, say Chesney and Citron (2019).

The qualitative analysis suggests that the mechanism will be essential for a better grasp of the wider political implications of deepfakes. A manufactured video doesn't have to be convincing to everyone to have political repercussions. It can, alternatively, breed a sense of cynicism generally, so that political entities can take advantage of uncertainty when faced with embarrassing evidence. This is an imbalance between misinformation and accountability; people may not only be less able to figure out what is false, but what is true.

Table 3. The Liar's Dividend and Evidentiary Trust

Mechanism	Qualitative interpretation	Consequence for politics
Existence of convincing deepfakes	Audiences become aware that audiovisual evidence can be fabricated	Authentic evidence becomes more contestable
Strategic denial	Actors can claim genuine evidence is synthetic	Accountability may be weakened
Generalized skepticism	Citizens become uncertain about competing claims	Trust in journalism and institutions may decline
Verification burden	Authenticity requires additional evidence and institutional confirmation	Political communication becomes slower and more complex
Epistemic fragmentation	Different groups may accept incompatible versions of reality	Polarization and political disagreement may intensify

Note. Based particularly on the conceptualization of the liar's dividend by Chesney and Citron (2019), supplemented by the broader literature on synthetic media and political trust.

The discovery has significant ramifications for journalism, democratic accountability and diplomacy. If when the authenticity is not easily established, then, the importance of the institutions with higher credibility and verification infrastructure increases. The question is not only one of cutting down on incorrect information, but one of enhancing the institutional ability to delineate and convey what is reasonably authentic.

Theme 4: Deepfakes, Diplomacy, and International Security

The fourth theme relates deepfakes to diplomatic communication, information warfare, foreign interference and international security. The ability to exchange diplomatic messages relies heavily on credible signals and accurate interpretation of official messages. Audio or video fabricated by these political leaders, government officials or military authorities may then raise doubts about intent, decision or action.

This is more important during periods of crisis when governments might only have one or two days to confirm information. An artificial rumour about military action, diplomacy, or emergency action may cause confusion until it is proven true. But the result doesn't necessarily create a situation of inter-state conflict. Instead, they point to a combination of factors—the speed, the uncertainty, the difficulty in attribution, and the political tension—that can raise the risk of miscalculation.

Table 4. Diplomatic and International-Security Implications

Domain	Potential deepfake application	Qualitative risk	Institutional requirement
Diplomatic signaling	Fabricated statements by political leaders	Misinterpretation of intentions	Rapid diplomatic verification
Crisis communication	False emergency announcements	Confusion during crises	Trusted official communication channels
Foreign interference	Synthetic political content targeting foreign audiences	Manipulation of public opinion	Cross-border information cooperation
Information warfare	Coordinated synthetic narratives	Strategic deception	Information-resilience capabilities
Military communication	Fabricated military statements or events	Miscalculation	Secure authentication and attribution

The damage to international public diplomacy, manipulated representation of states or leaders, transparent communication and verification.

The results corroborate the information-warfare view of information as an arena of strategic competition. The term deepfakes is not synonymous with the term 'warfare'; it is better seen as another weapon in the arsenal of influence operations. They are effective only when they are timed, distributed, consumed by the right audience, and only when they have the credibility of other sources of information.

This is also an interpretation that ties in with constructivist perspectives on international relations. Partly, international political behaviour is influenced by common meanings, identities, expectations, and interpretations. Disturbing processes of creating credible representations of events with synthetic media can have an impact on social construction of diplomatic reality.

Theme 5: Governance, verification and democratic resilience

The fifth theme is focused on the reactions to deepfakes. The analysis suggests that a single technological and/or regulatory solution is not enough. While some manipulated content can be detected using AI, detection may not work as well as the frequency of the evolution of generative models. The technologies that can support the notion of provenance and watermarking can give information on the origin of the content, but the degree to which they can work depends on the degree of adoption, interoperability and keeping of related metadata. While fact-checking can offer helpful context check, it may not be able to keep up with the pace of viral synthetic content.

The results provide evidence for a multi-layered resilience strategy, which includes technological, institutional, educational, regulatory and international solutions.

Table 5. Governance and Resilience Responses

Response mechanism	Potential contribution	Major limitation	Required complementary measure
AI detection	Identifies potentially manipulated content	Detection techniques can become outdated	Continuous technical development
Content provenance	Provides information about origin and modification	Depends on adoption and metadata integrity	Industry-wide interoperability
Watermarking	Signals AI-generated content	Can be removed or inconsistently implemented	Provenance and platform enforcement
Fact-checking	Provides contextual assessment	Often slower than viral dissemination	Rapid institutional communication
Media literacy	Improves citizens' critical evaluation	Effects vary across audiences	Continuous public education
Platform governance	Can reduce harmful amplification	Transparency and enforcement remain contested	Independent oversight
Election regulation	Establishes rules for political AI content	Enforcement across platforms and borders is difficult	Electoral and platform cooperation
International cooperation	Addresses transnational information operations	Different legal and political standards	Shared norms and verification protocols

The main conclusion is that governance needs to be concerned with the production of and the interpretation of. The methods of detection are primarily established in order to discover manipulated content, and a resilient communication also requires trusted institutions, media literacy, transparent political communication and credible means of authenticating official communication. International cooperation is particularly important as synthetic products can cross jurisdictions quickly, before many regulations.

Cross-Theme Synthesis

Overall, the five themes demonstrate the need to view deepfakes as a political challenge rather than one that is solely technical. The information landscape is embedded in electoral manipulation, diplomatic uncertainty, foreign interference, and international security threats, all of which are becoming issues of verifiable information as the issue of authenticity comes into play. The main analytical strength of this study is that it aims to establish epistemic uncertainty as the link that binds these domains.

The findings have shown that the biggest political danger of deepfakes may not be that they can mislead viewers to believe false information, but that they can also be believed to be a source of accurate information. Instead, in many situations, their value can be seen as a diminished trust in the information environment. If citizens can't tell if political evidence is real or fake, if journalists have to spend more to verify, if diplomats can't instantly verify the authenticity of the statements made by foreign leaders, if political actors can dismiss true evidence as false, the impact is far more than just on individual pieces of false evidence.

In this regard, a holistic approach to democratic resilience, strategic communications, information warfare, and responsible governance of AI must also incorporate deepfakes. Technological alarmism and institutional complacency has no place. Instead, complementary investments are needed in content provenance and authentication, in responsible platform governance, in media literacy, in electoral safeguards, in diplomatic verification mechanisms, in cybersecurity cooperation, and international

norms must be strengthened. Any policy measures that are effective are best implemented through a process of building trust in the public, enabling legitimate political expression and not unnecessarily limiting communication.

Conclusion

As the production and circulation of synthetic media become more affordable and available to an ever increasing number of people, the process of creating, circulating and assessing political information has been revolutionised. Deepfakes are not only a technological break-through in the area of AV manipulation, but a challenge to the credibility and politicisation of evidence in a world where the originality of the communication and its falsification can be even harder to tell apart. This study has shown that they do not necessarily have to sway a listener or viewer to accept or reject a fabricated audio or video message as true or false but are instead meant to generate epistemic uncertainty about political information in general. Deepfakes' impacts extend beyond just deceptive content, as Vaccari and Chadwick (2020) demonstrate by showing that those who come into contact with synthetic political media become more uncertain and less confident about the information sources.

The qualitative results also suggest that this uncertainty can have significant consequences for electoral integrity and democratic legitimacy. Deepfakes have the potential to be used for award impersonation, reputation attacks, fake political statements or deceptive campaign and marketing messages. Their influence should not, however, be considered as a given voting and outcomes factor. Rather, the democratic function might be expressed in disorientation, polarization, damage to reputation, and loss of faith in electoral institutions. In fact, the "liar's dividend" is closely related to this, as politicians can now label real information as fake, as they can now easily disseminate synthetic media (Chesney & Citron, 2019). So the issue is not just about the diffusion of false information, it is about a diminishment of common grounds for determining the true information.

The challenges also manifest themselves in the field of diplomacy, foreign interference, information warfare and international security. In some critical moments, there may be confusion since some important political or military statements might have been misinterpreted, and miscalculations are more likely to happen. In this respect, deepfakes introduce an additional layer of complexity to the international political context where timely and accurate attribution and credible signaling are among key factors in maintaining strategic stability. Based on the results, synthetic media is a phenomenon of a wider process of change in the global information environment in which the power can now be conveyed through the manipulation of perception, credibility and common political sense.

A multi-level governance solution is the answer to the above challenges. Governments, EMBs, diplomats, international organisations, technology providers, journalists, civil society and cybersecurity institutions, and election preparedness measures, must be strengthened in the areas of authentication, content provenance, institutional verification, digital literacy, platform accountability, responsible AI governance, and election preparedness measures. The problem, however, can't just be solved with technology. Technological solutions are still a complement to the other instruments to build resilience, such as the legal, institutional, professional journalism, democratic norms and international cooperation.

The field of generative AI is rapidly developing, the evidence from different countries is not always consistent, and there is little transparency about the use of the technologies, which adds to the methodological difficulties in establishing direct causal links between the use of deepfakes and political outcomes. Such limitations, however, are not an impediment to the underlying analytical thesis. Today, however, the new challenge in international politics is not only to distinguish between real and synthetic media, but also to keep institutions that can provide credibility to their claims in spite of the challenge. As synthetic political communication becomes increasingly sophisticated, the strength of global politics will be based not so much on the ability to eliminate the fabrication entirely, but rather on the institutional, technological and societal capabilities to verify the political reality, in a world of uncertainty.

References

- Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211–236. <https://doi.org/10.1257/jep.31.2.211>
- Barari, S., Lucas, C., Munger, K., & others. (2024). Human detection of political speech deepfakes across transcripts, audio, and video. *Nature Communications*, 15, 7629. <https://doi.org/10.1038/s41467-024-51998-z>
- Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, 33(2), 122–139. <https://doi.org/10.1177/0267323118760317>
- Bradshaw, S., & Howard, P. N. (2019). *The global disinformation order: 2019 global inventory of organised social media manipulation*. Oxford Internet Institute, University of Oxford.

- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>
- Chandra, B., Dunietz, J., Roberts, K., Lee, Y., Fontana, P., & Awad, G. (2024). Reducing risks posed by synthetic content: An overview of technical approaches to digital content transparency (NIST AI 100-4). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.AI.100-4>
- Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review*, 107(6), 1753–1820. <https://doi.org/10.15779/Z38RV0D15J>
- Clayton, K., Blair, S., Busam, J. A., Forstner, S., Gance, J., Green, G., Kawata, A., Kovvuri, A., Martin, J., Morgan, E., & others. (2020). Real solutions for fake news? Measuring the effectiveness of general warnings and fact-check tags in reducing belief in false stories on social media. *Political Behavior*, 42, 1073–1095. <https://doi.org/10.1007/s11109-019-09533-0>
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE.
- Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Kendeou, P., Vraga, E. K., & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, 1(1), 13–29. <https://doi.org/10.1038/s44159-021-00006-y>
- Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science*, 363(6425), 374–378. <https://doi.org/10.1126/science.aau2706>
- Hameleers, M., van der Meer, T. G. L. A., & Dobber, T. (2024). Distorting the truth versus blatant lies: The effects of different degrees of deception in domestic and foreign political deepfakes. *Computers in Human Behavior*, 152, 108096. <https://doi.org/10.1016/j.chb.2023.108096>
- International IDEA. (2024). Artificial intelligence for electoral management. International Institute for Democracy and Electoral Assistance.
- International IDEA. (2024). Protecting democratic elections through safeguarding information integrity. International Institute for Democracy and Electoral Assistance.
- Kietzmann, J., Lee, L. W., McCarthy, I. P., & Kietzmann, T. C. (2020). Deepfakes: Trick or treat? *Business Horizons*, 63(2), 135–146. <https://doi.org/10.1016/j.bushor.2019.11.006>
- Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, 359(6380), 1094–1096. <https://doi.org/10.1126/science.aao2998>
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE.
- Łabuz, M., & Nehring, C. (2024). On the way to deep fake democracy? Deep fakes in election campaigns in 2023. *European Political Science*, 23(4), 454–473. <https://doi.org/10.1057/s41304-024-00482-9>
- Marwick, A., & Lewis, R. (2017). *Media manipulation and disinformation online*. Data & Society Research Institute.
- Mirsky, Y., & Lee, W. (2021). The creation and detection of deepfakes: A survey. *ACM Computing Surveys*, 54(1), Article 7. <https://doi.org/10.1145/3425780>
- Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). Jossey-Bass.
- Nightingale, S. J., & Farid, H. (2022). AI-synthesized faces are indistinguishable from real faces and more trustworthy. *Proceedings of the National Academy of Sciences*, 119(8), e2120481119. <https://doi.org/10.1073/pnas.2120481119>
- Patton, M. Q. (2015). *Qualitative research & evaluation methods* (4th ed.). SAGE.
- Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*, 25(5), 388–402. <https://doi.org/10.1016/j.tics.2021.02.007>
- Rid, T. (2020). *Active measures: The secret history of disinformation and political warfare*. Farrar, Straus and Giroux.
- Roozenbeek, J., & van der Linden, S. (2019). Fake news game confers psychological resistance against online misinformation. *Palgrave Communications*, 5, Article 65. <https://doi.org/10.1057/s41599-019-0279-9>

- Shukla, A. K., & Tripathi, S. (2024). AI-generated misinformation in the election year 2024: Measures of European Union. *Frontiers in Political Science*, 6, 1451601. <https://doi.org/10.3389/fpos.2024.1451601>
- Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A., & Ortega-Garcia, J. (2020). DeepFakes and beyond: A survey of face manipulation and fake detection. *Information Fusion*, 64, 131–148. <https://doi.org/10.1016/j.inffus.2020.06.014>
- Tracy, S. J. (2010). Qualitative quality: Eight “big-tent” criteria for excellent qualitative research. *Qualitative Inquiry*, 16(10), 837–851. <https://doi.org/10.1177/1077800410383121>
- Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, 6(1), 1–13. <https://doi.org/10.1177/2056305120903408>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>
- Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review*, 9(11), 39–52. <https://doi.org/10.22215/timreview/1282>
- Woolley, S. C., & Howard, P. N. (Eds.). (2016). *Computational propaganda: Political parties, politicians, and political manipulation on social media*. Oxford University Press.
- World Economic Forum. (2024). *Global risks report 2024* (19th ed.). World Economic Forum.
- UNESCO. (2023). *Guidelines for the governance of digital platforms: Safeguarding freedom of expression and access to information through a multistakeholder approach*. United Nations Educational, Scientific and Cultural Organization.



2026 by the authors; Journal of The Kashmir Journal of Academic Research and Development. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).