Volume 05, Issue 01, 2026. pp. 83–96
Review Article
DOI: https://doi.org/10.62078/grks.2026.v05i01.007

GS WOW
Wisdom of Worthy Research Journal

Political Communication in Nepal: Artificial Intelligence, Deepfakes and Democratic Governance

Sunil Paudel
Department of Journalism and Mass Communication, Madan Bhandari Memorial College

Article Info.

Email
sunilpaudel.np@gmail.com

Article History
Received: May 12, 2026
Accepted: July 10, 2026

Cite
Paudel, S. (2026). Political communication in Nepal: Artificial intelligence, deepfakes and democratic governance. *GS WOW: Wisdom of Worthy Research Journal, 5*(1), 83–96. https://doi.org/10.62078/grks.2026.v05i01.007

Abstract

Generative tools like artificial intelligence (AI) can now create fake speeches, fake videos, and fake voices that look and sound real. This is changing the way political power is won, used, and questioned. Using constructivist approach of review and qualitative document analysis method this paper studies how this technology, especially AI, deepfakes, is affecting political communication and democratic governance, with special attention to Nepal. In September 2025, a government ban on social media platforms triggered the Gen Z protests, the deadliest unrest in the country's democratic history that led to the resignation of the government. Six months later, as Nepal prepared for the March 2026 House of Representatives election, the Election Commission reported widespread deepfake videos, hate speech, and AI-generated images aimed at candidates and voters. This paper reviews some global academic literature and connects it with Nepal's National Artificial Intelligence Policy 2025, the Gen Z protests, and the 2026 election. The findings show that deepfakes damage trust in leaders, media, and elections, and this damage can be more harmful than the fake content itself. Nepal needs stronger digital literacy programs, faster fact-checking systems, clear election laws on AI content. This paper is meant to contribute by providing suggestions to policymakers, researchers, and citizens for protecting democracy in the age of artificial intelligence.

Keywords: artificial intelligence, deepfakes, democratic governance, political communication

Introduction

Political communication is the way governments, leaders, parties, and citizens share information, arguments, and promises with each other. For most of history, this communication moved through speeches, newspapers, radio, and television. Today, it moves mostly through the internet, and increasingly, through artificial intelligence. Artificial intelligence, or AI, means computer systems that can learn, reason, and create content that looks like it was made by a human being (Ministry of Communications and Information Technology [MoCIT], 2025). In the last five years, AI tools have become cheap, fast, and easy to use. Anyone with a phone and an internet connection can now create a fake photo, a fake video, or a fake voice recording of a real person. These fake creations are called deepfakes.

Deepfakes are synthetic media, meaning audio, video, or images, that are built or changed using artificial intelligence to make a real person appear to say or do something they never said or

did (Chesney & Citron, 2019). At first, deepfakes were mostly used for entertainment or harmless jokes. But very quickly, political actors around the world began to use them to attack rivals, confuse voters, and escape responsibility for their own actions (Pawelec, 2022). Scholars now warn that deepfakes threaten the basic working of democracy: free elections, informed citizens, a trusted media, and honest public debate (Woolley & Howard, 2018; Tenove, 2020).

Nepal gives this global problem a very local and very urgent face. Nepal is a young democracy. It ended 239 years of monarchy in 2008 and started democratic republic system. It has been struggling since a long history with weak governments, frequent leadership changes, and public anger over corruption (Teekah, 2026). In September 2025, this anger exploded. The government tried to force 26 social media platforms, including Facebook, YouTube, Instagram, and WhatsApp, to register with the state or be blocked (Shamin, 2025). Many young Nepalis, especially the group known as Generation Z, saw this as an attack on free speech and a way to hide corruption, symbolised by viral "Nepo Kid" posts showing the children of politicians living rich lifestyles while ordinary youth struggled to find jobs (Pokharel, 2025). The protests that followed killed dozens of people, burned government buildings including the parliament, and forced the Prime Minister to resign (Pokharel, 2025; Participedia, 2025).

Only months later, as Nepal prepared for a new House of Representatives election on March 5, 2026, a second and different problem appeared: a flood of AI-made disinformation. The Election Commission reported more than 600 cases of hate speech and deepfake content sent to authorities, including fake videos showing candidates insulting each other and fake images of massive rallies that never happened (Saxena, 2026; Malay Mail, 2026). A manipulated photograph falsely labelled a candidate as a foreign agent (Gahatraj, 2026). A deepfake audio clip of a senior political figure spread widely before fact-checkers could catch it (Gahatraj, 2026). These events show that Nepal is not a side note in the global story of AI and democracy. It is now one of the clearest real-world examples of how fast this threat can arrive, even in a country with limited digital infrastructure.

At the same time, Nepal has tried to respond. In August 2025, the Government of Nepal approved its first National Artificial Intelligence Policy, called National Artificial Intelligent (AI) Policy, (2025). The policy promises ethical AI use, data protection, and a new AI Regulation Council (MoCIT, 2025). But experts already ask whether this policy is strong enough, funded enough, and fast enough to manage the risks already visible on the ground (Shah, 2026).

This paper studies this whole picture together: the global rise of AI in political communication, the specific danger of deepfakes, and what this means for democratic governance, using Nepal as the central case. The goal is not only to describe the problem, but to give useful, evidence-based ideas for policymakers, election bodies, journalists, and citizens.

Democratic governance is fundamentally dependent on public trust. Citizens must have confidence that the information they receive is accurate, that electoral processes are conducted fairly, and that political leaders can be held accountable for their statements and actions. However, the rapid advancement of artificial intelligence (AI) has transformed the way information is created, disseminated, and consumed, presenting unprecedented opportunities as well as significant societal challenges (Mishra & Mishra, 2025; Ananda & Mishra, 2025; Vasuki et al., 2026). While AI has enhanced decision-making, organizational resilience, and digital transformation across sectors, its misuse, particularly through deepfake technology, poses a serious threat to democratic institutions and public trust (Vasuki et al., 2026; Mishra & Mishra, 2025).

Deepfake technologies threaten democracy in two important ways. First, AI-generated audio,

video, and image content can deliberately mislead citizens by creating false narratives about political candidates, political parties, or public events. Second, the existence of deepfake technology enables individuals to reject authentic evidence by claiming that genuine content has been artificially generated, thereby weakening political accountability and public confidence in digital evidence. Chesney and Citron (2019) describe this phenomenon as the "liar's dividend," whereby the ability to deny authentic digital evidence becomes a powerful political strategy. As AI capabilities continue to advance, distinguishing authentic information from manipulated content is becoming increasingly difficult, thereby amplifying the risks associated with misinformation and disinformation (Mishra & Mishra, 2024; Mishra & Mishra, 2025).

Although developed democracies have begun addressing these challenges through comprehensive regulatory initiatives such as the European Union's AI Act and sophisticated fact-checking mechanisms, developing countries such as Nepal face substantially greater challenges. Nepal continues to experience uneven digital literacy, limited technical and institutional capacity for detecting AI-generated misinformation, and a recently introduced legal framework that has yet to be adequately tested (MoCIT, 2025; Ghimire, 2026). Previous studies have shown that successful AI implementation depends not only on technological advancement but also on institutional preparedness, governance structures, and public awareness, all of which remain underdeveloped in many emerging economies (Mishra, 2019; Ananda & Mishra, 2025).

Recent political developments in Nepal further illustrate the seriousness of this challenge. The 2025 Gen Z protests demonstrated how rapidly online information—whether accurate or misleading—could mobilize citizens and influence political behaviour through digital platforms (Participedia, 2025). Likewise, the 2026 general election highlighted the growing potential for AI-generated misinformation and deepfakes to manipulate public opinion, damage political reputations, and undermine confidence in democratic elections (Saxena, 2026). These developments indicate that Nepal's democratic institutions are confronting rapidly evolving AI-enabled threats while policy implementation, institutional capacity, technological preparedness, and public awareness continue to lag behind technological advancement (Vasuki et al., 2026).

Although the Government of Nepal introduced the National Artificial Intelligence Policy 2082 (2025) to promote responsible AI development and governance, limited academic research has evaluated its effectiveness in addressing AI-driven political disinformation and deepfake technologies. Existing literature by Mishra and colleagues primarily emphasizes AI adoption, digital transformation, educational innovation, employee engagement, and organizational performance (Mishra & Mishra, 2025; Ananda & Mishra, 2025; Vasuki et al., 2026), while empirical evidence concerning AI's implications for democratic governance and electoral integrity in Nepal remains scarce. Moreover, international scholarship on deepfakes and democracy has largely focused on developed democracies in North America and Europe, leaving South Asian contexts relatively underexplored.

Therefore, the central problem addressed by this study is the widening gap between the rapid advancement of AI and deepfake technologies and Nepal's comparatively limited legal, institutional, technological, and societal capacity to safeguard democratic governance from AI-driven misinformation and political manipulation. By examining global theoretical perspectives alongside Nepal's experiences during the 2025 Gen Z protests and the 2026 general election, and by evaluating the implementation of the National Artificial Intelligence Policy 2082, this study seeks to generate evidence-based recommendations for strengthening democratic resilience, promoting responsible AI governance, and enhancing public trust in Nepal's evolving digital democracy.

Literature Review

Political Communication in the Age of AI

Artificial intelligence is not one single tool. It includes machine learning systems that sort and recommend content, chatbots that can write political messages, and generative systems that can create images, video, and audio from a short text instruction (MoCIT, 2025). In political communication, all of these are now in daily use. Parties use AI to target voters with personalised messages. Campaigns use AI to write speeches and social media posts quickly. Government offices explore AI to speed up public services. Alongside these useful applications, AI has also produced a darker set of tools, most notably deepfakes and automated "bot" accounts that flood social media with fake engagement (Woolley & Howard, 2018).

A large academic literature has grown around what researchers call computational propaganda: the use of algorithms, automated accounts, and human-guided coordination to spread misleading political information over social media (Woolley & Howard, 2018). Woolley and Howard's edited collection studied nine countries, both democratic and authoritarian, and found that political bots and fake accounts were used to make small groups look like large, popular movements, to attack opponents, and to manipulate what appeared to be "trending" online (Woolley & Howard, 2018). Earlier country-based research, such as work on the 2016 Brexit referendum, found that a small number of highly active accounts produced a very large share of all political messages on a hashtag, showing how automation can distort the appearance of public opinion (Woolley & Howard, 2018). This body of research is important because it shows that the danger from AI in politics is not only about fake video content. It is also about fake scale: making a small, coordinated effort look like a spontaneous, mass movement.

Growth in Political Use of Deepfakes

A deepfake is created using a class of AI models that can study many images or recordings of a real person and then generate new, synthetic content showing that person saying or doing something new (Chesney & Citron, 2019). The word combines "deep learning," the AI method used, and "fake." Early deepfakes, mostly made around 2017 and 2018, were used for non-consensual pornography, but the technology's political use grew quickly after that (Chesney and Citron, 2019).

Chesney and Citron's widely cited 2019 article in the California Law Review remains one of the most complete academic studies of deepfakes' danger to privacy, democracy, and national security. They describe several harms: deepfakes can be used to fabricate evidence against a political opponent, to fake a candidate's concession speech, or to create panic before an election (Chesney & Citron, 2019). Their most important and lasting contribution, however, is the concept of the "liar's dividend." As people become more aware that video and audio can be faked, some individuals accused of real wrongdoing use this awareness to falsely claim that authentic, damaging evidence against them is a deepfake (Chesney & Citron, 2019). This means that simply knowing deepfakes exist can protect wrongdoers, even when the actual fake technology was never used. This idea has since been applied to many real cases, from politicians denying real recordings to authoritarian governments dismissing verified evidence of human rights abuses as fake (Pawelec, 2022).

Building on this legal analysis, Pawelec (2022) offers one of the few studies that connects deepfakes directly to democratic theory. Using Warren's (2017) "problem-based" theory of democracy, Pawelec argues that healthy democratic systems must deliver three things: empowered inclusion, meaning that everyone affected by a decision can take part in shaping it; collective agenda and will formation, meaning that citizens can debate and form shared views through fair deliberation; and collective decision-making, meaning that decisions taken have real legitimacy (Warren, 2017, as cited in Pawelec, 2022). Pawelec's review of more than 300 sources found that deepfakes damage all three functions. Deepfake pornography, mostly

aimed at women and dissidents, pushes them out of public debate, damaging inclusion. Election-related deepfakes and the liar's dividend damage the quality and honesty of public deliberation. And the resulting drop in trust damages the legitimacy of whatever decision is finally made, including election results (Pawelec, 2022). Importantly, Pawelec's review found that deepfakes had not yet decisively changed the outcome of any major election as of 2022, but that the fear of deepfakes alone was already damaging trust in elections, which she argues may be the technology's biggest political effect so far (Pawelec, 2022).

More recent scoping reviews confirm that deepfake research is growing quickly across many fields, including political science, computer science, and communication studies, but that most studies still focus on Western democracies, leaving countries like Nepal understudied (Momeni, 2025). A systematic review published in 2026 found that political deepfakes can be as persuasive and credible to viewers as authentic media, which is a serious warning sign for any country without strong fact-checking systems (Barari et al., 2025, as cited in ScienceDirect systematic review, 2026).

Global Responses for AI Governance and Regulation

Governments around the world have started responding to these risks, though responses vary widely. The European Union adopted the AI Act, the first broad AI law in the world, which sorts AI systems by risk level and bans the most dangerous uses, such as government social scoring, while placing strict rules on high-risk uses like hiring or medical decisions (Shah, 2026). India launched the IndiaAI Mission in 2024 with a large budget to build computing power, train AI researchers, and support AI start-ups (Shah, 2026). These examples show two different regulatory paths: strict, rights-based regulation in Europe, and investment-driven growth strategy in India. Smaller and developing democracies, including Nepal, must decide which mix of these two paths fits their own needs, resources, and risks.

AI Policy in Nepal

Nepal's own engagement with AI governance is very recent. The Ministry of Communication and Information Technology formed a technical committee in January 2025 to study AI and prepare Nepal's first concept paper on the subject (Rising Nepal, 2024). This concept paper recommended that Nepal create a national AI policy, a national strategy, and a legal framework for data protection, while comparing practices from the European Union, India, China, and other countries (Rising Nepal, 2024). This work led, in August 2025, to Cabinet approval of the National Artificial Intelligence Policy 2082, Nepal's first dedicated AI policy.

The policy's stated vision is to build "a human-centred, ethical, and prosperous Nepal through Artificial Intelligence technology" (National AI Policy 2082, as cited in Digital Watch Observatory, 2025). It builds on earlier foundations, including the Electronic Transactions Act 2006, the ICT Policy 2015, the Digital Nepal Framework 2019, and the National Cybersecurity Policy 2023 (Digital Watch Observatory, 2025). The policy creates a National AI Centre and an AI Regulation Council, sets a target of training 5,000 AI professionals within five years, and calls for AI excellence centres in all seven provinces (Digital Watch Observatory, 2025). It explicitly names deepfakes, data privacy, bias, and job losses as risks the government must manage.

Even so, Nepali commentators have already raised doubts. An opinion piece in Shah (2026) argues that Nepal's policy is "a true step in the right direction" but warns that without a dedicated budget, a truly independent AI Regulation Council, and serious public education, the policy risks staying "on paper" while real-world harms, like the deepfakes seen in the 2026 election, continue to spread faster than the government can respond. This gap between policy design and policy delivery is a central theme of the present paper.

A second, distinct but connected, strand of Nepali digital politics is the 2025 Gen Z protest

movement. Multiple international sources agree on the basic sequence of events: on September 4, 2025, the government ordered the shutdown of 26 social media platforms after they failed to meet a new registration requirement tied to digital taxation rules (Shamim, 2025). Officially, the government said the rule was about registering a local contact and grievance officer to fight cybercrime and fake accounts (Shamim, 2025). Critics, however, argued that the real trigger was a viral online trend exposing the lavish lifestyles of politicians' children, nicknamed "Nepo Kids". Mass protests began on September 8, 2025, led largely by Nepal's Generation Z, roughly ages 13 to 28 (Shamim, 2025). Security forces used tear gas, water cannons, and live ammunition; at least 19 to 22 people died in the first days alone, with the death toll later reported near 74 by mid-September (Harvey et al., 2025; Kathmandu Post, as cited in Wikipedia, 2026). Parliament, the Supreme Court, and other government buildings were set on fire, and Prime Minister KP Sharma Oli resigned (Harvey et al., 2025). An interim government led by former Chief Justice Sushila Karki was then formed to guide the country toward new elections (Ghimire, 2026).

Scholars studying this event highlight two important points for this paper. First, digital platforms were not only a battleground but also the organising tool of the movement itself: protestors used Discord servers and Instagram channels, run partly through the group Hami Nepal, to share safety information and mobilise supporters, even using VPNs and printed QR codes to work around the government's own ban (Participedia, 2025). Second, the uprising echoed earlier youth-led movements in Sri Lanka (2022) and Bangladesh (2024), suggesting a wider regional pattern of young, digitally connected citizens rising against corruption and perceived censorship (Teekah, 2026)). For this paper, the Gen Z protests are important because they prove that Nepali citizens, especially the young, are both highly capable digital organisers and highly vulnerable to state efforts at digital control, setting the stage for the AI-driven disinformation problems that followed in the 2026 election.

Research Gap

Taken together, the global literature gives strong theoretical tools, especially Chesney and Citron's (2019) liar's dividend and Pawelec's (2022) democratic-functions framework, and the Nepal-specific sources give rich, very recent factual material. However, almost no existing academic work directly connects these two bodies of knowledge. Most deepfake and democracy research still centres on the United States, the European Union, and large Asian democracies like India. Nepal, despite now offering one of the clearest live examples of AI-driven election disinformation anywhere in the world, remains largely outside the academic conversation. This paper responds directly to that gap.

Theoretical Framework

This paper uses Warren's (2017) problem-based theory of democracy, as extended by Pawelec (2022) to the study of deepfakes, as its main theoretical lens. This framework does not ask whether a specific formal rule or institution exists. Instead, it asks whether a political system actually delivers three functions that most democratic theorists agree are essential.

The first function is empowered inclusion: everyone who is affected by a political decision should be able to take part in shaping it, and should have real power, not just a formal right, to do so (Warren, 2017). The second function is collective agenda and will formation: citizens must be able to debate, argue, and form shared views and priorities through fair and honest public discussion (Warren, 2017). The third function is collective decision-making: when a decision is finally made, for example, through an election, it must be seen as legitimate by the people it affects (Warren, 2017).

Applying this framework to Nepal, this paper treats deepfakes, computational propaganda, and the government's social media ban all as forces that can either support or damage these three functions.

For example, the 2025 social media ban blocked platforms that many citizens used for political voice, directly threatening empowered inclusion (Teekah, 2026)). The flood of deepfakes during the 2026 election threatened collective agenda and will formation, because voters could no longer be sure which videos, images, or claims were real (Saxena, 2026). And both events, together, threaten the legitimacy of collective decision-making, since an election held in an atmosphere of heavy disinformation and recent political violence carries a heavier burden of proof to be seen as fair and trustworthy.

This paper also uses the concept of computational propaganda (Woolley & Howard, 2018) as a secondary lens, to understand not only individual fake videos but also the coordinated, large-scale patterns of manipulation, such as networks of fake accounts or AI-generated images of fake rallies, seen in the Nepal case (Malay Mail, 2026; Ojha, 2026).

Research Objectives

To examine the impact of artificial intelligence (AI) and deepfake technologies on political communication and democratic governance, with particular emphasis on AI-driven disinformation in Nepal's 2025 Gen Z protests and 2026 general election, and to evaluate the effectiveness of the National Artificial Intelligence Policy 2082 in addressing these challenges while proposing practical measures to safeguard democratic governance.

Methodology

Research Design

This paper uses a qualitative, document-based research design. Because the subject combines global theory with a fast-moving, real-time national case, a qualitative approach is more suitable than a numerical survey, since it allows the deep reading and comparison of many different kinds of sources: peer-reviewed journal articles, government policy documents, and verified news reporting (Pawelec, 2022, uses a similar mixed-document method).

Data Sources

Three types of sources were used. First, peer-reviewed academic literature on deepfakes, computational propaganda, and democratic theory, including Chesney and Citron (2019), Pawelec (2022), Woolley and Howard (2018), and Warren (2017). Second, official government and policy documents, including Nepal's National Artificial Intelligence Policy 2082 and related government statements. Third, verified reporting from established news organisations, including Reuters-linked wire coverage, CNN, Al Jazeera, and specialised Nepal-based and South Asia-focused outlets, describing the 2025 Gen Z protests and the 2026 general election.

The paper uses thematic, qualitative document analysis. Sources were read and grouped into themes matching the theoretical framework: threats to empowered inclusion, threats to collective agenda and will formation, and threats to collective decision-making. Within each theme, global evidence and Nepal-specific evidence were compared side by side, allowing the paper to test whether patterns seen in Western academic literature also appear in the Nepali case, and where the Nepali case shows something new.

Case Selection

Nepal was chosen as the central case for three reasons. First, it experienced two major, connected digital-political events within roughly six months: the September 2025 Gen Z protests and the March 2026 general election. Second, both events are extremely recent and well documented by multiple independent international sources, giving strong evidence quality. Third, Nepal represents a type of country, a small, developing democracy with uneven digital literacy and a brand-new AI policy, that is currently missing from most deepfake and democracy research, which usually studies large, wealthy democracies.

Results and Discussion

Social Media Ban and The 2025 Gen Z Protests

On September 4, 2025, Nepal's government ordered the Nepal Telecommunications Authority to block 26 social media platforms, including Facebook, X, YouTube, WhatsApp, LinkedIn, Reddit, Signal, and Snapchat, after the companies failed to register with the Ministry of Communication and Information Technology by a set deadline. The registration rule required each platform to name a local contact person, a grievance officer, and a self-regulation officer (Shamim, 2025). The government linked this requirement to a new digital services tax and stricter value-added tax collection from foreign technology companies.

Public reaction, however, focused less on tax policy and more on freedom of expression and corruption. In the days before the ban, a wave of "Nepo Kid" content had gone viral, showing the children of senior politicians travelling abroad, wearing luxury brands, and living comfortably while Nepal's youth unemployment stood at roughly 20 percent (Harvey et al., 2025; Shamim, 2025). About 90 percent of Nepal's population uses the internet, and a large share of the population, more than 7 percent, lives and works abroad, depending on platforms like Messenger and Viber to stay in contact with family (Shamim, 2025). The sudden ban, therefore, struck at both political anger and everyday communication needs at the same time.

Demonstrations began on September 8, 2025, gathering first at Maitighar Mandala, a well-known public square in Kathmandu, before moving toward the federal parliament (Shamim, 2025). Security forces responded with tear gas, water cannons, and, according to multiple reports, live ammunition (Pokharel, 2025). At least 19 people died on the first day in Kathmandu alone, with additional deaths in other cities; later reporting put the overall death toll from the unrest at approximately 74, described by one regional outlet as the highest single-day protest death toll in Nepal's history (Harvey et al., 2025). On September 9, protestors entered the parliament complex, and by the next day, the parliament building, the Supreme Court, and other government offices, along with the private residence of the Prime Minister, had been set on fire (Harvey et al., 2025). Prime Minister KP Sharma Oli resigned, and the government reversed the social media ban within days of the first protest (Pokharel, 2025).

In the aftermath, an interim government led by former Chief Justice Sushila Karki was formed, with a mandate to prepare the country for a new general election (Ghimire, 2026). International bodies, including the United Nations human rights office, expressed shock at the death toll and called for a transparent investigation into the use of force by security personnel (Harvey et al., 2025).

From a political communication standpoint, three points stand out. First, the protest movement was itself a highly organised digital communication campaign, coordinated through Discord servers and Instagram channels, some connected to the youth-led group Hami Nepal, even while official platforms were blocked, showing the resilience of digital organising against state control (Participedia, 2025). Second, the trigger for the unrest was not artificial intelligence directly, but it shared a key feature with AI-driven disinformation problems: both involve a breakdown of trust between citizens and the state over what information is real, fair, and honestly presented. Third, the scale of citizen anger released by the ban shows how sensitive digital communication rights have become in Nepal, which is exactly the kind of high-tension environment in which deepfakes and disinformation can do the most damage, a link that became clear only months later, in the 2026 election.

The 2026 General Election and the Rapid Growth of Deepfakes

Following the interim government's mandate, Nepal held a House of Representatives election on March 5, 2026 (Gahatraj, 2026). In the weeks leading up to the vote, election observers, journalists, and the Election Commission

itself documented a sharp rise in AI-generated disinformation, described by one wire report as an "onslaught" (Saxena, 2026).

The Election Commission's information officer, Suman Ghimire, stated publicly that there was "widespread use of hate speech and deepfake content," including videos made with easily available AI tools that falsely showed candidates insulting opponents or using obscene language (Saxena, 2026; Malay Mail, 2026). More than 600 such cases were referred to authorities, with about 150 handled directly by police (Saxena, 2026). In one case, police detained a pro-royalist supporter, Durga Prasai, over social media posts allegedly intended to intimidate voters (Malay Mail, 2026). Loyalists connected to the recently removed government also shared AI-generated images falsely presented as drone photographs of a rally attended by more than 500,000 supporters; these images were then shared further by senior party leaders, showing how AI-generated content can move from anonymous accounts into mainstream political messaging within hours (Ojha, 2026).

A Center for Media Research Nepal survey found that more than 95 percent of online users in Nepal had encountered false information of some kind, at a time when 73 percent of Nepalis use smartphones and over 37 percent have internet access, based on the country's most recent national census figures (Ghimire, 2026). Researcher Hikmat Acharya of the fact-checking platform TechPana described spending months verifying suspicious content, sometimes relying on AI detection tools such as Hive Moderation and Google SynthID (Gahatraj, 2026). One case that stood out involved a deepfake audio clip attributed to a senior political figure, Swarnim Wagle, which spread widely before it could be confirmed as fake (Gahatraj, 2026).

Individual candidates were also directly targeted. In late January 2026, a manipulated photograph of Sobita Gautam, a candidate for the Rastriya Swatantra Party in Chitwan Constituency No. 3, spread across Nepali social media with a false label calling her an "American agent." The image had actually been taken from a public youth policy event in Kathmandu held in July 2025, co-organised by the United States Embassy Youth Council and the Youth Innovation Lab, but was stripped of its real context and reused to suggest foreign interference (Gahatraj, 2026). A separate viral video of another candidate, Balendra Shah, showing him pushing his own party workers, was first shared approvingly by supporters as a sign of strong leadership before becoming part of a wider dispute over its authenticity and framing (myRepublica, 2026).

In response, the Election Commission invited representatives from major platforms, including TikTok, to discuss enforcement, and the government directed law enforcement agencies and the Cyber Bureau to act quickly against confirmed misinformation (Gahatraj, 2026; myRepublica, 2026). However, commentators noted that with the election only weeks away at the time these steps were announced, there was little room left for a nationwide digital literacy campaign, and that corrections from fact-checkers rarely travelled as far or as fast as the original false content (myRepublica, 2026).

Nepal's National Artificial Intelligence Policy 2082

Nepal's policy response to these risks did not begin with the 2026 election. It began earlier, in January 2024, when the Ministry of Communication and Information Technology formed a five-member technical committee to study global AI practice and prepare Nepal's first concept paper on AI (Rising Nepal, 2024). That concept paper, submitted after nearly five months of study, recommended that the government create a national AI policy, a national strategy, and a legal framework for data protection, while addressing cybersecurity and privacy risks (Rising Nepal, 2024). It compared AI strategies from the European Union, the United Nations, India, China, and several other countries.

This process led to Cabinet approval of the National Artificial Intelligence Policy 2082 in

August 2025. The policy's vision is "to build a human-centred, ethical, and prosperous Nepal through Artificial Intelligence technology," and its mission is to use AI to support Nepal's socio-economic development across health, education, agriculture, industry, finance, and public administration (Digital Watch Observatory, 2025). The policy builds on earlier digital governance instruments, including the Electronic Transactions Act 2006, the ICT Policy 2015, the Digital Nepal Framework 2019, and the National Cybersecurity Policy 2023 (Digital Watch Observatory, 2025).

On institutions, the policy creates a National AI Centre to act as a secretariat and coordinating body, and an AI Regulation Council that must meet at least twice a year and can invite outside officials or experts as needed (MoCIT, 2025). On goals, it sets specific targets: ranking Nepal among the top 50 countries on a Government AI Readiness Index, training at least 5,000 skilled AI professionals within five years, and establishing AI excellence centres in all seven provinces (Digital Watch Observatory, 2025). On risk, the policy explicitly names data privacy, algorithmic bias, deepfakes, and job displacement as challenges the government must actively manage, and calls for a formal AI governance framework aligned with international standards, including UNESCO's principles for ethical AI (Digital Watch Observatory, 2025).

Nepali commentary on the policy has been cautiously positive but clearly worried about implementation. An opinion piece in Shah (2026) describes the policy as "a true step in the right direction" but warns that Nepal risks repeating a common pattern seen in other developing countries: publishing an ambitious policy without matching it with a dedicated budget, a genuinely independent regulatory body, or serious public education efforts. The same piece points out that neighbouring India has committed more than 10,371 crore rupees over five years to its own AI mission, raising the question of whether Nepal's policy, however well written, can be implemented without comparable financial commitment (Shah, 2026). The piece further argues that the AI Regulation Council must be staffed with genuinely independent experts rather than political appointees if it is to be trusted (Shah, 2026).

This paper's own analysis, using the three-function democratic framework from Section 7, finds that the policy is strong on paper regarding long-term goals such as skills training and sector-by-sector AI adoption, but comparatively thin on the more urgent, election-specific problem of deepfakes and disinformation. The policy names deepfakes as a risk but, based on available public documentation, does not yet set out detailed, binding rules for political AI content during election periods, an area that the 2026 election showed to be immediately and practically urgent, arguably more urgent in the short term than many of the policy's longer-term sectoral goals.

Bringing together the global literature and the Nepal case material, several clear patterns emerge.

First, Nepal confirms Chesney and Citron's (2019) warning about the liar's dividend, and extends it into a live democratic setting. During the 2026 election, once the public understood that AI could fabricate convincing videos and audio, it became harder for voters, journalists, and even the Election Commission to be fully confident about any single piece of content, whether fake or genuine. This matches Pawelec's (2022) finding that the fear of deepfakes can damage trust in elections just as much as, or more than, confirmed deepfakes themselves.

Second, Nepal shows a pattern of computational propaganda, in Woolley and Howard's (2018) sense, appearing alongside individual deepfakes. The AI-generated images of a fake, massive rally (Ojha, 2026) were not aimed at a single target the way a deepfake video of one candidate is. Instead, they were aimed at manufacturing an illusion of scale and popularity, echoing Woolley and Howard's (2018) finding that automated and AI-assisted content is often used to make a political message look far more popular than it really is.

Third, and specific to Nepal, the paper finds a direct link between the 2025 Gen Z protests and the 2026 election disinformation problem that has not yet been studied together in existing literature. The social media ban of September 2025 was fought, in large part, in the name of digital rights and freedom of expression (Teekah, 2026)). Yet only months later, the same digital space that citizens fought to protect became the main channel for AI-generated attacks on the very election meant to renew the country's democratic mandate (Saxena, 2026). This suggests that simply defending open digital platforms, while necessary, is not sufficient on its own to protect democratic governance; open platforms must be paired with strong content verification systems, or they risk being used against the same democratic goals citizens fought to defend.

Fourth, applying the three-function democratic theory from Section 7, the case shows damage across all three areas at once, which is unusual and concerning. Empowered inclusion was threatened twice, first when the social media ban blocked citizens' political voice, and again when deepfakes falsely labelled a candidate a foreign agent, potentially discouraging voters from supporting her, or discouraging other candidates and citizens from entering public debate for fear of similar attacks (Gahatraj, 2026). Collective agenda and will formation was damaged by the sheer volume of disinformation, with more than 95 percent of online Nepalis reporting exposure to false content (Ghimire, 2026), making honest public debate about real policy issues far harder to sustain. Collective decision-making, meaning the legitimacy of the election result itself, was placed under strain by the combination of recent political violence, a very short window for fact-checking before voting day, and doubts about whether viral content, rally sizes, or candidate statements could be trusted (myRepublica, 2026).

Fifth, the case also shows some encouraging signs. Nepali fact-checkers and platforms like TechPana actively used AI detection tools to counter AI-generated fakes, showing that the same technology causing the problem can also be part of the solution, if properly resourced (Gahatraj, 2026). The Election Commission did take direct action, including inviting platform representatives for discussion and referring hundreds of cases to police, showing institutional willingness to respond, even if capacity remained limited relative to the scale of the problem (Gahatraj, 2026; Saxena, 2026).

Overall, the Nepal case supports the global academic warning that deepfakes and AI-driven disinformation are a serious threat to democratic governance, while also showing that this threat interacts strongly with a country's specific political history, in Nepal's case, deep public anger over corruption, a very young population, high smartphone use, and a very new and still-developing policy framework.

Limitations of the Study

This study has several limitations that should guide how its findings are read. First, because the 2026 Nepal election and its aftermath were recent and, in some respects, still unfolding at the time of writing, some sources are early news reports rather than fully completed academic research, and details such as final vote counts, court cases, or long-term effects on public trust may become clearer only later. Second, the study relies on a qualitative, document-based method rather than direct surveys or interviews with Nepali voters, election officials, or fact-checkers, which means it cannot measure, in numerical terms, exactly how much specific deepfakes changed individual voting decisions. Third, most of the deep theoretical literature used in this paper, particularly Chesney and Citron (2019) and Pawelec (2022), was originally built around cases from the United States, Europe, and other established democracies; while this paper argues the theory applies well to Nepal, some adjustment may be needed to fully capture unique local factors such as caste, ethnicity, religion, and rural-urban digital divides, which fell outside this paper's main focus. Fourth, information about Nepal's National AI Policy 2082 is based on publicly available summaries and commentary rather than

a full legal-technical reading of every clause of the policy document, so some specific implementation details may not be fully captured here. Future research using direct surveys, interviews with Nepali election officials and fact-checkers, and a full legal analysis of the National AI Policy 2082 would help confirm and extend the findings of this paper.

Conclusion

Artificial intelligence is changing political communication faster than most democratic institutions can adapt. Deepfakes and AI-driven disinformation attack the basic trust that democracy needs to function: trust in what citizens see and hear, trust in fair elections, and trust in accountable leaders. Global academic research, especially Chesney and Citron's (2019) idea of the liar's dividend and Pawelec's (2022) democratic-functions framework, shows how deep this danger runs, even before a single fake video changes an election result.

Nepal's recent experience makes this danger concrete rather than theoretical. The September 2025 Gen Z protests showed how quickly digital communication conflicts can escalate into national political crisis, at the cost of dozens of lives and the fall of a government. The March 2026 general election showed how the same digital space can become a battlefield of AI-generated deepfakes, fake rally images, and manipulated audio, all arriving faster than fact-checkers, police, or the country's brand-new National Artificial Intelligence Policy 2082 could fully manage.

This paper has argued that Nepal's policy response, while a genuine and important first step, is not yet matched to the speed and scale of the threat, particularly around elections. Protecting democratic governance in the age of artificial intelligence will require Nepal, and countries like it, to move from general, long-term AI policy goals toward specific, well-funded, election-ready systems of verification, transparency, and public education. The stakes are high: as this paper has shown, the cost of falling behind is not only fake videos online, but real damage to public trust, real political violence, and, ultimately, real damage to the legitimacy of democratic government itself.

Based on the analysis above, this paper offers the following recommendations for Nepal, while noting that similar lessons may apply to other developing democracies facing early-stage AI governance challenges.

Recommendation

- Nepal's Election Commission should create clear, specific rules for AI-generated political content during election periods, including mandatory labelling of AI-made images, video, or audio used in campaigns, going beyond the general risk statements currently found in the National AI Policy 2082.
- The government should give the National AI Centre and the AI Regulation Council a dedicated, protected budget and staff them with independent technical experts rather than political appointees, directly addressing the concern raised by Nepali commentators that the policy could remain "on paper" without proper funding.
- A standing, well-funded partnership should be built between the Election Commission, the Cyber Bureau, independent fact-checking organisations such as TechPana, and major social media platforms, so that verification tools such as Hive Moderation and Google SynthID can be used quickly during high-risk periods like elections, not only after content has already gone viral.
- Digital and media literacy education should be added to school curricula and public awareness campaigns well before election periods begin, rather than attempted in the final weeks before voting, when, as this paper's case study shows, there is little time left for such efforts to work.

- Any future response to harmful online content, including social media regulation, should be designed transparently and separately from unrelated goals such as tax collection, to avoid the kind of public backlash and loss of trust seen during the September 2025 social media ban.
- Nepal should study and adapt relevant parts of international models, including the European Union's risk-based AI Act and India's AI Mission's investment approach, rather than choosing only one path, since Nepal's challenges include both immediate risks needing strict rules and long-term development needs requiring investment.

References

Ananda, N., & Mishra, A. K. (2025). AI architecture for educational transformation in higher education institutions. *Poornaprajna International Journal of Management, Education & Social Science* (*PIJMESS*), *2*(2), 58–73. https://doi.org/10.5281/zenodo.16976456

Barari, S., Lucas, C., & Munger, K. (2025). Political deepfakes are as credible as other fake media and (sometimes) real media. *The Journal of Politics*, *87*(2), 510–526. https://doi.org/10.1086/732990

Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review, 107*, 1753–1819. https://doi.org/10.15779/Z38RV0D15J

Digital Watch Observatory. (2025). *Nepal's national AI policy 2082 (2025)*.

Gahatraj, D. (2026, March 4). As Nepal goes to the polls, deepfakes and AI manipulation undermine democracy. *South Asia Monitor*.

Ghimire, S. (2026, January 30). Surge of disinformation clouds Nepal's upcoming federal election as AI and deepfakes add challenges. *NepYork*.

Harvey, L., Pokharel, S., Mitra, E., & Iyer, A. S. (2025, September 9). A social media ban, corruption and 'Nepo Kids:' What we know about the deadly protests that ousted Nepal's leader. *Cable News Network* (CNN)

Malay Mail. (2026, March 3). Nepal election: How AI disinformation is threatening the vote that changed everything.

Mishra, A. K. (2019). Implementation status of value management in project management practice in Nepal. *International Journal of Management Studies, 6*(1), 92–108. https://doi.org/10.18843/ijms/v6i1/12

Mishra, A. K., & Mishra, S. (2025). *Artificial and emotional intelligence for employee engagement with e-human resource management*. Intellectuals' Book Palace. https://doi.org/10.5281/zenodo.14810072

Mishra, S., & Mishra, A. K. (2024). AI influencing factors among students. *Rabi Sangyan, 1*, 1–8. https://doi.org/10.3126/rs.v1i1.74673

Ministry of Communications and Information Technology. (2025). *National artificial intelligence (AI) policy, 2025*.

Momeni, M. (2025). Artificial intelligence and political deepfakes: Shaping citizen perceptions through misinformation. *Journal of Creative Communications*, *20*(1), 41–56. https://doi.org/10.1177/09732586241277335

myRepublica. (2026, April 4). Deepfakes and democracy.

Ojha, A. (2026, March 4). AI disinformation turns Nepal polls into 'digital battleground'. *The Japan Times*.

Participedia. (2025). 2025 Nepal Gen Z protests.

Pawelec, M. (2022). Deepfakes and democracy (theory): How synthetic audio-visual media for disinformation and hate speech threaten core democratic functions. *Digital Society, 1*(2), Article 19. https://doi.org/10.1007/s44206-022-00010-6

Pokharel, S. (2025, September 8). Nepal lifts social media ban that sparked deadly Gen Z protests. *Cable News Network* (*CNN*).

Saxena, V. (2026, March 3). Nepal facing AI disinformation 'onslaught' ahead of key polls. *Asia Financial*.

Shah, S. (2026, April 14). Nepal's AI policy exists: But is Nepal ready? The Himalayan Times.

Shamim, S., (2025, September 8). Nepal Gen Z protests amid social media ban, clashes kill 19: All to know. *Al Jazeera*.

Teekah, E. (2025, October 7). 2025 Nepalese Gen Z protests (Editorial). *Britannica*.

Tenove, C. (2020). Protecting democracy from disinformation: Normative threats and policy responses. *The International Journal of Press/Politics, 25*(3), 517–537. https://doi.org/10.1177/1940161220918740

The Rising Nepal. (2024, July 4). Nepal prepares its first concept paper on AI use.

Vasuki, M., Mishra, A. K., Kumar, A. D., & Celestin, M. (2026). Artificial intelligence and supply chain resilience: Implications for organizational performance. *Journal of Advanced Research in Quality Control and Management, 11*(2), 24–34. https://doi.org/10.24321/2582.3280.202610

Warren, M. E. (2017). A problem-based approach to democratic theory. *American Political Science Review, 111*(1), 39–53. https://doi.org/10.1017/S0003055416000605

Woolley, S. C., & Howard, P. N. (Eds.). (2018). *Computational propaganda: Political parties, politicians, and political manipulation on social media*. Oxford University Press. https://doi.org/10.1093/oso/9780190931407.001.0001