Between Progress and Peril: The Role of Artificial Intelligence (AI) in Shaping Modern Political Communication

Radoslav Baltezarević[1]* , *Vladimir Lović[2]* , *Ivana Baltezarević[3]*

[1]Institute of International Politics and Economics, Belgrade, Serbia, e-mail: radoslav@diplomacy.bg.ac.rs
[2]International center for economics and public policy - ICEPP, Belgrade, Serbia, e-mail: vladalovic@yahoo.com
[3]Institute for Political Studies, Belgrade, Serbia, e-mail: ivana.baltezarevic@ips.ac.rs

Abstract: Since artificial intelligence (AI) has been integrated into our digital communication landscape, there have been major changes in how political campaigns are strategically designed and how public opinion is influenced. With the help of machine learning (ML), generative models (GM) and natural language processing (NLP), AI tools have introduced new opportunities for political engagement. Today, thanks to AI-driven data analytics, we can micro-target voters based on their psychographic profiles and adapt political messages with incredible precision. On the other hand, generative AI technologies are increasingly used to spread false information or to imitate political endorsements, which has a great impact on public opinion. The dissemination of such content can greatly reinforce ideological prejudices and contribute to social divisions. This paper draws on recent empirical research and case studies to illustrate how AI-generated disinformation campaigns can affect electoral processes and undermine trust in democratic institutions. Various examples, such as the use of bots to control social media to deepfake content impersonating political figures, show that ethical, technological and legal safeguards are urgently needed. Furthermore, this paper supports an approach to AI governance that strikes a balance between promoting innovation and reducing harm. This implies the development of tools for AI detection, transparency measures and cooperation between sectors in order to promote responsibility and integrity of information. Greater digital literacy among citizens and proactive policy responses will be necessary in the near future to ensure the resilience of democratic systems due to the increasingly rapid development of AI technology.

Keywords: Artificial Intelligence (AI), Political Communication, Generative AI, Deepfakes, Disinformation, Large Language Models (LLMs).

Introduction

According to Fetzer and Fetzer (1990), artificial intelligence (AI) is the ability of computers to imitate human intelligence and to think and learn in the same ways as human beings. The impact of AI, due to its rapid development, is becoming greater in numerous fields, especially in politics. The rise of digital technology has brought a host of highly complex challenges to the sharing, reliability and consumption of political information. This shift prompts a careful examination of the role of AI in political communication and highlights the need for strategies that promote greater transparency and, in our digital age, foster informed public discussions (Bareis and Katzenbach, 2022).

The acceptance of cultural values and public opinion are greatly influenced by the media, many of which are deceptively promoted as independent ideas (Baltezarević et al., 2014). However, the rise of new technologies is increasingly suppressing traditional cultural values and narratives, often replacing them with algorithmically generated content (Baltezarević et al., 2019). AI has revolutionized the way we analyze extensive data sets, including demographic, behavioral, and psychological data derived from individuals' online behaviors, as well as the growing issue of deepfakes (Chester and Montgomery, 2017). These tools are often used to create persuasive political messages (Carr, 2011).The emergence of "deepfakes," that is, realistic fake videos produced through AI-based facial manipulation, has raised social alarms due to their

ability to imitate a person's actions or speech with minimal signs of manipulation (Chawla, 2019).

Alexander Nix (the former CEO of Cambridge Analytica) mentioned that by understanding the psychological characteristics of target audiences, political communications can be adjusted to connect with specific traits, be it emotional, rational, or fear-driven (Mermoud, 2017). Large language models (LLMs) are typically fine-tuned to boost user engagement by spotting patterns that create emotionally resonant and engaging content (Ethayarajh et al., 2024). A focus on engagement can often amplify controversial narratives, as people tend to connect with content that is consistent with their beliefs (Brady et al., 2017).

More than half of Americans voiced serious concerns about the spread of AI-generated political propaganda, according to a 2023 survey, while 23% were only moderately concerned (Statista Research Department, 2024). These worries are not unfounded, as there have already been several reports of AI being abused in political communication. For instance, the founder of the investigative website Bellingcat, Eliot Higgins, created fake photos of Donald Trump's arrest using the Midjourney AI platform, and the German political party AfD used AI-generated graphics to inflame anti-refugee sentiment (Matheis, 2023).

According to a recent survey, social media is one of the main sources of news for 56% of respondents, but almost 70% of them believe that it is also the largest source of false information, raising serious questions about its reliability (Weforum, 2024). This dynamic, with the advent of generative artificial intelligence (GenAI), only adds to the complexity. GenAI systems can both create and identify fake content (Loth et al., 2024). For example, OpenAI's GPT models are able to produce content that closely resembles the style, tone, and structure of reliable news sources, making it difficult to distinguish between authentic journalism and made-up stories (Brown et al., 2020).

To tackle this challenge, technology and human oversight must collaborate. Large language models (LLMs) can indeed spread false information, but they also have the potential to help us fight against it (assuming we implement human intervention effectively). It's vital for users, content moderators, and news organizations to take proactive measures to curb the dissemination of misleading information. Reporting questionable content can enhance AI detection systems and help quickly identify harmful or deceptive information (Virginia Tech, 2024).

On a broader level, AI has significant implications for democratic processes. The ability of AI to create misleading information, influence voter behavior, and potentially undermine election integrity is a major concern (Coeckelbergh, 2022). Yu, 2024 in their study highlights the necessity of finding a balance between robust regulatory measures, technological advancements, and ethical oversight to mitigate these risks. With the potential to distort political realities and threaten democratic institutions, deepfakes and AI-generated misinformation are among the most pressing threats to the integrity of elections.

Ultimately, it's essential to reflect on the ethical implications and the potential of AI technologies as they advance. Since AI is still in the early days of development, continuous research, development, and critical evaluation are key to solving current issues and shaping the future of this technology in a way that bolsters democratic processes instead of undermining them (Baltezarević and Baltezarević, 2024).

In order to explore this dynamic more thoroughly, this paper is guided by the following research questions:

1. In what way does artificial intelligence (AI) influence and change strategies in the field of political communication in modern democratic societies?
2. How does AI-powered voter profiling influence the shaping of public opinion?
3. What ethical and legal issues emerge from the political use of deepfakes and AI-generated misinformation?
4. How do AI algorithms lead to political polarization on social media by creating filter bubbles and echo chambers?
5. What legal and technological measures can be implemented to minimize the dangers of AI misuse while still upholding free speech and democratic values?

The Role of AI in Political Communication, Campaign Strategy, and Public Opinion Manipulation

AI is defined as any computer or algorithm that can observe its surroundings, learn from them, and make smart decisions based on that knowledge. This definition is quite expansive and includes a variety of technologies, but machine learning (ML) techniques are currently some of the most popular approaches (Samoili et al., 2020).

AI is dramatically changing the way politicians connect with the public and shape opinions. Its abil-

ity to analyze complex data, process information, and adjust communication strategies based on user preferences makes it a game-changer in modern politics (Crawford, 2021). One of the key applications of AI in this field is data-driven voter profiling, where potential voters are linked to personality types through AI-powered analytics. When we classify personality types using psychographic and demographic factors, we can make informed predictions about how people will respond to different stimuli. This skill gives campaigns a crucial advantage, helping them to sway swing voters and inspire their target audience to participate in elections (Wakefield, 2019).

To make sense of the huge amounts of data flowing from social media and online platforms, politicians are increasingly turning to AI-based analytics to get a clearer picture of what voters really think and feel (Battista and Uva, 2023). This growing reliance on AI tools is enhancing election campaigns and communication strategies, making them more targeted and effective (Alvarez et al., 2023). AI also allows for real-time monitoring of public sentiment on political issues, debates, and candidates. Campaigns can leverage this data to identify potential voters and tailor personalized ads that either encourage or discourage them from voting. Consequently, political leaders can adjust their positions and communication strategies to keep up with the constantly shifting landscape of online public opinion (Lutkevich and Hildreth, 2022).

AI is revolutionizing campaign management by not only boosting financial contributions but also enhancing how messages are crafted. A striking example of this is the "Vote Leave" campaign during the Brexit referendum, which cleverly used A/B testing to tailor their messages for specific demographic segments. Campaign director Dominic Cummings explained that the catchphrase "Let's take back control" was deeply rooted in a careful analysis of public opinion on the European Union (EU). Through their iterative testing, they found that the word "back" in the slogan sparked anger by triggering loss aversion, the psychological tendency to prefer avoiding losses over acquiring equivalent gains, especially when it comes to control (Schneider, 2017).

To more effectively reach and engage their target voters, candidates and their campaign staff can harness the predictive power of AI by combining data from a variety of sources (Crilley, 2018). This capability allows for the development of highly customized political messages that resonate with voters' unique needs and preferences, enhancing the campaign's overall impact (Nunziata, 2021). The use of AI is growing rapidly, especially for real-time fact-checking and analyzing public reactions to political events. For example, Microsoft's AffectiveSpotlight employs head movements and facial expressions to assess viewers' emotional responses, allowing presenters to tweak their delivery style (Murali et al., 2021).

Political debates can benefit greatly from AudienceView, which uses large language models (LLMs) to classify audience feedback, thus directly helping journalists better understand public opinion (Brannon et al., 2024). In addition to this system, Factiverse and Full Fact are real-time fact-checking websites that have introduced AI systems capable of capturing, evaluating, and analyzing political commentary in real time (Corney et al., 2024). A billion web pages may seem impossible to scan. However, Factiverse is making it feasible for both companies and individuals to extract the information they require into essential, easily assimilated insights using intelligent prioritization and robust crawling tools (Factiverse, 2025). Full Fact is a British nonprofit that verifies and corrects news reports and statements that make the rounds on social media (Dudfield, 2025). Full Fact AI uses cutting-edge artificial intelligence to enable fact checkers, journalists, researchers, and communicators to recognise, validate, and refute misleading content (Fullfact, 2025).

AI's impact stretches far beyond just campaigning; it's also leaving its mark on governance. One of the standout advantages is its knack for sifting through massive amounts of data to predict the outcomes of political decisions. AI has also changed how people communicate with their representatives, introducing fascinating new features like chatbots and virtual assistants on company websites and AI-powered social media interactions. This has opened up more direct communication channels between elected officials and the people they represent (Viudes, 2023). Still, there are some risks to consider. For instance, during the 2025 protests in Los Angeles, AI chatbots such as Grok and ChatGPT were found to spread false information because they misinterpreted images and lacked the necessary context (Gilbert, 2025).

AI chatbots can provide voters with tailored insights about candidates and policies, which can boost participation and support informed decision-making (Political Communication, 2023). However, the rise of AI-generated content, including deepfakes and synthetic media, poses a serious risk to the integrity of political communication (Thornhill, 2024). These tools can be misused to spread false information, influence public opinion, and erode trust in democratic institutions, particularly in nations with lower media literacy levels (Funk et al., 2023). To truly harness the benefits of AI, it's essential for governments, civil

society, and tech companies to collaborate and create strategies that are specifically designed for different contexts (Hagerty and Rubinov, 2019).

We tend to underestimate the importance of improving our digital literacy, especially given the threats posed by deepfakes and AI-generated disinformation. The speed at which synthetic media is evolving is outpacing our technology and legal systems. Encouraging media literacy and critical thinking skills has become a more effective and scalable way to address these problems. Recent research shows that digital literacy programs can successfully reduce people's vulnerability to false information by improving their ability to spot content that is manipulative and better understand how certain narratives are promoted by algorithmic systems (Guess et al., 2020). It is also crucial that both governments and digital platforms work together to promote public education through accessible initiatives, such as media warnings, verified fact-checking labels, and very clear explanations of algorithms (Roozenbeek and van der Linden, 2020). Even the most sophisticated detection technologies can become useless if there is no widespread digital literacy, as end users remain vulnerable to emotionally charged synthetic content (Sustainability Directory, 2025).

AI-driven bots have become quite common on social media platforms. Typically powered by ML algorithms, these bots can imitate human behavior at an impressive speed while still appearing genuine. They use natural pauses and responses to create the illusion that a human is behind the messages (Bessi and Ferrara, 2016). Any government or political party with enough financial backing can deploy a whole army of social bots to influence public discussions on social media, removing the need for specialized technical skills and resources (Ferrara et al., 2016).

While these technologies can enhance social interactions and deepen our understanding of one another, they also come with risks like deceit, manipulation, and the spread of misinformation (Gallo et al., 2022). The influence of social bots is amplified by how easy they are to create and manage. Social bots often take on tasks such as researching hashtags and keywords, posting content, responding to user interactions, following users interested in specific topics, and gathering opinions on online discussions (Ferrara et al., 2016). The misinformation generated by AI, especially the biased or fake news produced by advanced text generation systems, could seriously undermine political engagement and pose a threat to democratic processes (Klinger et al., 2023).

We can theoretically distinguish between "good" and "bad" social bots. Bad (malicious) bots usually spread harmful links or misleading stories, while good bots share posts that offer valuable information. Bad bots often create fake accounts that look almost identical to real users, making it tough to differentiate them from trustworthy content. As noted by BSI (2025), these bots typically engage in coordinated "fake news" efforts that seek to shape public opinion. Pamment et al. (2018) highlight that this dishonest strategy can significantly impact political backing and alter our perception of reality in multiple ways.

The growing presence of bots is only widening the gap in society, creating confusion, and shaking public trust in information. This makes it increasingly tough to tell what's real and what's not (Bradshaw and Howard, 2018). For the first time in over ten years, automated traffic has overtaken human internet usage, making up almost 51% of all web traffic, as highlighted in the 2025 Imperva Bad Bot Report. The surge is mainly driven by large language models (LLMs) and the swift rollout of AI technology (Chang, 2025).

AI is revolutionizing fields such as speech and sentiment analysis, as well as content creation and distribution. To grasp the emotional tone and reactions of the public, political campaigns can use sophisticated algorithms that analyze speeches and communications, helping them detect how audiences are feeling (Khare, 2023). Sentiment analysis is a vital part of natural language processing (NLP) and categorizes emotional content in text as neutral, negative, or positive (based on subjective data) (Rajashekhargouda, 2022). This analytical approach improves business intelligence and provides measurable insights that can improve strategic decisions (Kumar and Garg, 2020). Sentiment analysis models can reveal the polarity of opinions, the topics being discussed, and the individuals behind them (Obot et al., 2025).

AI is also influencing social media influencer marketing strategies outside of the political campaign space. Technologies created by a number of advertising agencies assess an influencer's "brand safety" and forecast whether or not they will participate in political discourse. A marketing company called Captiv8, which works with companies like Kraft Heinz, for example, recently introduced an AI-powered tool that analyses social media users' online mentions to determine how likely they are to discuss elections or other politically sensitive topics. In this system, an "A" means you should exercise extreme caution, while a "C" suggests a profile that's generally safe. Influencers receive these letter ratings based on their writing, comments, and how they're covered in the media. These ratings consider delicate matters like hate speech,

violence, sexual content, and divisive social issues (Maheshwari, 2024).

Artificial Deception: The Crisis of Democratic Integrity, Deepfakes, and Disinformation

The rise of AI brings with it some crucial ethical dilemmas, particularly around the accessibility and transparency of information (Nida-Rumelin and Weidenfeld, 2019). With AI technologies impacting every facet of how we produce and share information, these challenges are becoming more pressing in our rapidly changing digital world. One of the most concerning issues we are facing is the rise in deepfakes. This not only threatens the integrity of our political systems but also puts our society's trust and stability at risk (Westerlund, 2019).

Deepfakes are made by combining deep learning techniques with altered content. This technology creates a deceptive reality that can be difficult to detect with just a glance, allowing a person's face to convincingly express emotions or say things they've never actually uttered (Korshunov, 2018). Thorough training, the generator (Generative Adversarial Networks - GANs), enhances its ability to create more realistic fake images, driven by the competitive dynamics of the system (Shen, 2018). This has resulted to deepfakes becoming increasingly intricate and more difficult to spot.

A diverse array of groups, including governments, political activists, criminals, dubious individuals, conspiracy theorists, and automated bots, are leveraging deepfakes to achieve their goals. According to Zannettou (2019), these objectives can differ greatly, ranging from influencing public sentiment and inciting social chaos to seeking financial gain and demonstrating ideological loyalty. In the political arena, deepfakes are particularly worrisome. They can produce fake videos of people appearing to say or do things they never actually did. This can confuse voters and might even change the course of elections. Such manipulation is a serious threat to democracy, as it prevents individuals from making well-informed choices (Cheguri, 2023).

We must also be aware of the ways these technologies can be misused, such as in financial fraud, scams, hoaxes, fake news, non-consensual pornography, extortion, harassment, bullying, electoral interference, and the spread of disinformation. However, although deepfakes are often viewed negatively, they can actually be quite innovative and beneficial. They have the potential to vividly reenact historical events, enhancing the authenticity of films or serving as useful educational resources (Cruci, 2024). With the rapid advancement of deepfake creation methods and the development of AI detection tools to counter these risks, it's more important than ever for people to learn how to identify misleading media (Appel and Prietzel, 2022). The very AI technologies that make it possible to produce convincing deepfakes also power the tools used to identify them, resulting in a dynamic and ever-changing battleground between creators and defenders. However, this leads to an ongoing technological arms race.

The public's trust in reliable sources of information and organizations has really taken a hit due to these changes. Deepfakes, being trickier to spot than traditional fake news, only add to the cybersecurity challenges that both individuals and organizations face. Plus, since deepfakes blur the lines between reality and fiction, they often have a stronger impact on shaping public opinion (Hu et al., 2022). For instance, Google Trends data reveals that searches for "free voice cloning software" skyrocketed by an astonishing 120% from July 2023 to 2024 (Cruz, 2024), highlighting how accessible technology has made it for even amateurs to create altered audio recordings.

Deepfakes are captivating and novel, which helps them spread false information at lightning speed. Studies indicate that the top 1% of rumors usually don't get more than 1,000 views on platforms like Twitter (now X). In fact, even true news stories tend to have a smaller reach, illustrating how misinformation can travel faster and further than the truth (Vosoughi et al., 2018). To make things even trickier, the technology behind deepfakes is advancing at a pace that outstrips our current detection methods (Surfshark, 2025).

It's difficult to overlook the impact that AI and deepfakes are having on elections. A 2024 survey found that more than 75% of people globally are anxious about AI's potential impact on future elections. The United States and Singapore are particularly concerned, with 72% and 83% of respondents, respectively, voicing their worries (Petrosyan, 2025). Furthermore, large language models (LLMs) and AI tools like Midjourney, Google's Gemini, and OpenAI's ChatGPT have made it easier to normalize deepfakes across multiple platforms (Zandt, 2024).

According to data collected since 2017, 31% of deepfake cases have involved fraud, with particular focus areas being celebrities, and political figures. The targeting landscape shows that 35% of incidents were aimed at politicians or celebrities, while a significant 65% impacted the general public. Of these in-

stances, 27% were politically provocative, and 25% contained explicit content. US President Donald Trump was the most affected, being the target of 18% of the deepfakes involving politicians, which amounts to 25 incidents. Joe Biden faced 20 deepfakes, including voice robocalls that were manipulated and often linked to election matters. Additionally, well-known politicians like Kamala Harris and Volodymyr Zelenskyy also experienced multiple attacks (Surfshark, 2025).

Additionally, large companies are becoming more susceptible to crimes made possible by deepfakes. For example, using a deepfake impersonation of the CFO of the British engineering company Arup, fraudsters were able to successfully transfer $25 million to Hong Kong bank accounts in 2024. The scam was confirmed by the deceitful CFO and team members who participated in video calls (Noto, 2024). Another significant case occurred in 2019, when fraudsters used a deepfake voice of the president of a UK energy company to siphon off €220,000 (Somers, 2020).

A low-tech deepfake video that purported to show American politician Nancy Pelosi intoxicated went viral in 2019, deceiving many Facebook users with almost 2.5 million views (CBS News, 2019). Globally, political deepfakes have become more complex. In early 2024, robocalls featuring a deepfake voice of President Biden were sent to thousands of voters in New Hampshire, discouraging them from voting in the primary election. The fact that these audio recordings were created in less than 20 minutes for just $1 each shows how readily and affordably such false information may propagate (Seitz-Wald, 2024). During the 2023 annual news conference with Russian President Vladimir Putin, a student from St. Petersburg caught everyone's attention by asking a question. What made this moment particularly striking was the use of deepfake technology. The voice and appearance of this individual were actually created using a deepfake AI avatar of Putin himself (NBC News, 2023).

European laws could serve as a useful blueprint for tackling the issue of false information generated by AI. The EU's Digital Services Act requires tech platforms to evaluate the risks their products pose to society, particularly concerning democracy and elections. Moreover, these platforms are required to provide relevant data to independent experts to help evaluate their impacts (Hetrick, 2024). As policymakers around the world tackle the rapidly changing landscape of AI technologies, these regulatory frameworks are working to find a balance between the risks and rewards associated with AI.

On a global scale, we're starting to see legislative measures take shape in response to these challenges. To boost transparency, the European Union's AI Act of 2024 requires that AI-generated content be clearly labelled. At the same time, several states in the US have enacted laws that make it illegal to create and distribute harmful deepfakes, particularly those intended to influence elections or spread misinformation (Kumar, 2025).

Regulating AI in political communication raises a number of legal concerns, particularly with regard to platform liability, foreign jurisdiction, and the tension between free speech and censorship. The European Union's AI Act and Digital Services Act (DSA) aim to regulate AI use and hold online platforms accountable (European Parliament, 2022). These laws prohibit the use of manipulative AI and also require platforms to eliminate illegal or harmful content, like hate speech and misinformation, by specific deadlines (European Commission, 2021). Although these strategies are designed to protect users, critics warn that they might lead to the over-removal of lawful content, which could violate free speech rights (Lazaro Cabrera, 2024).

With the rise of deepfakes, misinformation, often referred to as "fake news", has become alarmingly prevalent on social media in recent years. This trend has had a profound effect on our economy, politics, democracy, and society as a whole (Burkhardt, 2017). Fake news is often fueled by financial or political motives aimed at influencing public opinion on divisive topics, and it typically leads to serious real-world repercussions (Barclay, 2018). Major digital platforms often come under criticism for not being clear about how they make their decisions. They have a significant duty to manage content and reduce the spread of harmful or misleading information (Gorwa, 2019). This lack of transparency can confuse users, as they wonder why certain posts are taken down while other harmful content stays up, making it difficult to hold anyone responsible (Roberts, 2019).

The lack of transparency in algorithmic content curation and moderation policies raises some serious concerns about inconsistent enforcement and potential bias, especially during politically charged moments (Citron and Pasquale, 2019). While a few platforms have made strides to bring in third-party auditors and share transparency reports, these efforts often fall short of fully explaining the processes behind content moderation and appeals (Gillespie, 2018). This gap highlights important ethical and legal dilemmas about how to balance preventing harm with protecting free speech, as well as who really holds the responsibility

for overseeing the vast expanse of online content (Suzor, 2019).

AI algorithms tend to favor content that ramps up user engagement, often leading to the amplification of divisive or sensational topics, which only deepens the political divide. By shutting users off from a variety of perspectives and reinforcing their existing biases, these AI-powered recommendation systems create echo chambers and filter bubbles (Islam et al., 2024). Echo chambers are particularly common on social media platforms, where users are mostly exposed to information that aligns with their beliefs. This not only reinforces their deeply held convictions but also makes it increasingly difficult for them to accept differing opinions (Jiang et al., 2021). Although the phrase "echo chamber" does not appear often in political science literature, it is frequently used in the context of digital media studies. It illustrates how individuals with like-minded opinions come together online, getting constant reinforcement that can skew their understanding of reality and stifle genuine dialogue (Parry, 2006). Algorithmically personalized environments, often referred to as "filter bubbles," present significant challenges to democracy. These bubbles create separate realities instead of fostering the shared truths necessary for informed participation in a democratic society, as they limit users' exposure to diverse opinions. Such personalization reinforces pre-existing biases and desires by protecting people from information that challenges their beliefs, much like subtle propaganda (Populismstudies, 2018).

Global data highlights troubling trends in how people perceive news bias and the spread of misleading information, contributing to what's been dubbed the "infodemic." Roughly 60% of people worldwide say that most individuals only embrace information that backs up their own beliefs. This is particularly noticeable in countries such as the US (68%), Turkey (69%), Serbia (70%), and Peru (71%). What's intriguing is that while 65% of respondents feel that people are actively hunting for opinions that reinforce their views, only 34% admit to feeling stuck in an informational bubble (Konopliov, 2024). The power of this issue is highlighted by the fact that over 70% of Europeans often encounter fake news (Watson, 2024).

The influence of fake news touches every corner of society, affecting how people, groups, and governments respond to the false information that spreads like wildfire on social media. Much of this misinformation is crafted to target specific demographics, aiming to stir up conflict and strengthen ideological backing (Tandoc et al., 2018). In the US, political memes and viral videos are everywhere, often featuring altered images, clips taken out of context, and even portraits created by AI. A couple of striking examples include videos of Taylor Swift and Vice President Kamala Harris impersonating Donald Trump and seemingly endorsing him (Bond, 2024).

AI has really changed the game when it comes to creating fake news. Large Language Models (LLMs) can churn out a ton of readable and coherent content, thanks to their training on massive datasets. Plus, with rapid video generators like Sora, which can create detailed, Hollywood-style fake clips, the spread and impact of misinformation have only grown (Virginia Tech, 2024). To tackle this problem, AI models like Grover have been developed specifically to spot AI-generated fake news, and this approach has shown to be quite effective (Zellers et al., 2019). It analyses the text and structure of articles to find bias, inaccurate information, or other warning indicators using a combination of generation and detection strategies. For separating machine-generated news from human-written news, Grover claims an accuracy of above 92% (Gillham, 2025).

Discussion

The rapid advancement of artificial intelligence (AI) is changing politics and fundamentally altering how democracies function in the modern world. With tools like generative models, deepfake, and innovative content curation algorithms, we could see a surge in political participation, quicker communication, and more tailored interactions between citizens and their leaders. As campaigns get better at fine-tuning their messages, we could see a boost in voter turnout and overall democratic participation. Ideally, these innovations could revitalize democratic processes by making political conversations more accessible, dynamic, and aligned with the unique concerns of citizens.

The exciting potential of technology comes with some serious and complex risks that could undermine social cohesion and the integrity of our democracy. The negative aspects of this technological shift are starkly illustrated by deepfakes and AI-generated disinformation. Unlike simple text-based lies, deepfakes create realistic videos of real individuals making false or misleading statements. These can be

weaponized to attack opponents, suppress voter turnout, or increase social tensions, potentially destabilizing entire electoral processes. The political implications are severe. This kind of manipulation poses a significant threat to democracy, especially given how quickly and easily it can spread, as evidenced by deepfake robocalls aimed at voters or viral videos that misrepresent politicians.

Beyond just deepfakes, the polarization and division in public discourse are getting worse thanks to the widespread use of AI-driven content recommendation algorithms on social media. These algorithms tend to create echo chambers and filter bubbles by prioritizing content that boosts user engagement, which often means sensational, emotionally charged, or extreme ideological content. When users are cut off from opposing viewpoints, it only reinforces confirmation bias and intensifies political tribalism. This lack of common ground undermines the democratic dialogue and compromise we need, as it solidifies partisan rifts and diminishes faith in democratic institutions. Moreover, the worldwide spread of false information, often referred to as an "infodemic," adds to the challenge by sowing doubt about which news and information we can actually trust.

Furthermore, there's a concerted effort to regulate AI and deepfakes. The European Union is taking the lead in this area, working on comprehensive regulatory frameworks that impose responsibility, transparency, and content moderation on digital platforms. Key examples include the AI Act and the Digital Services Act (DSA). These laws require the identification of AI-generated content, mandate risk assessments, and ensure that harmful content is removed quickly, which are all crucial steps toward protecting democratic discourse. Despite the complex challenges at play, it's hard to say just how effective these restrictions will be. To begin with, politicians often struggle to keep pace with the rapid evolution of technology. On top of that, the global nature of the internet complicates jurisdiction enforcement, and the ongoing debates about free speech, censorship, and platform responsibility create tough moral dilemmas. Critics warn that overly broad regulations might hinder free expression, while too little oversight could allow misinformation to thrive.

On a more profound level, however, the phenomenon The Washington Post refers to as the "liar's dividend" is the increasingly prevalent practice of politicians claiming that compromising statements or videos were produced artificially (either by deepfake technology or other AI tools) when there is no supporting evidence (Verma and De Vynck, 2024). This approach sets a concerning precedent by leveraging uncertainty around the accuracy of digital content to dodge accountability. It muddles the line between manipulation and the actual truth. This trend sparks major concerns regarding democracy and social responsibility, as it diminishes the trust that is so important for meaningful discussions in the public sphere. Even when the information is accurate, people struggle to make informed choices if everything is labelled as potentially fake. This also gives those in power the ability to alter the truth without facing any consequences.

Keeping up with technological advancements is crucial in the battle against AI-driven misinformation. AI detection tools are getting better at spotting and flagging fake media. We really need to boost public awareness about AI-generated content and foster critical thinking to help people resist manipulation. Additionally, by putting transparency measures in place, like better content moderation disclosures, we can build more public trust and accountability.

The way societies tackle the ethical, legal, and technological challenges posed by AI will ultimately shape its impact on political communication and the integrity of democracy. We need to encourage collaborative governance frameworks that bring together governments, tech companies, academia, civil society, and the public to set guidelines and standards for ethical AI use. To find the right balance between fostering innovation and upholding essential democratic values like accountability, transparency, freedom of expression, and human rights, we need a strategy that involves multiple stakeholders. Rather than diminishing democratic participation, AI has the potential to enhance it, provided we integrate ethical considerations and maintain strict oversight in its development and application.

Conclusion

In conclusion, we must acknowledge AI's impressive potential to transform political communication and improve democratic engagement. However, we also need to be mindful of the serious threats posed by deepfakes and AI-driven misinformation campaigns to the future of this technology. It's essential that we take prompt action to combat the erosion of trust in information, the increasing social divides, and the

flaws in our electoral systems. Strong laws, advanced technology, public awareness, and global cooperation are vital for democracies to mitigate these risks and safeguard the integrity of their political institutions. The decisions we make today about AI regulation will either uphold the foundational principles of democracy or allow for manipulative technology that diminishes informed public engagement.

It's crucial for future research to dive deeper into interdisciplinary solutions that blend ethical AI design, media literacy education, and inclusive governance frameworks. Scholars, educators, policymakers, and tech developers need to join forces to create evidence-based strategies that strengthen civic resilience and safeguard democratic discourse. By nurturing a culture of transparency, critical thinking, and accountability in technology, we can ensure that AI serves as a tool for empowering democracy instead of causing disruption. The relationship between AI and politics is quite complex, and it requires our continuous joint efforts. We need to work together to nurture the intellectual and democratic development of future generations.

Acknowledgement

Funding

Conflict of interests

The authors declare no conflict of interest.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

Institutional Review Board Statement

Not applicable.

Author Contributions

Conceptualization, B.R., V.L. and B.I.; methodology, B.R.; formal analysis, B.R. and B.I.; writing—original draft preparation, B.R. and V.L.; writing—review and editing, B.R. and B.I. All authors have read and agreed to the published version of the manuscript.

References

Alvarez, R. M., Eberhardt, F., & Linegar, M. (2023, July). Generative AI and the Future of Elections (CSSPP White Paper). *California Institute of Technology Center for Science, Society, and Public Policy*. https://lindeinstitute.caltech.edu/documents/25475/CSSPP_white_paper.pdf

Appel, M., & Prietzel, F. (2022). The detection of political deepfakes. *Journal of Computer-Mediated Communication, 27*(4), zmac008. http://dx.doi.org/10.1093/jcmc/zmac008

Baltezarević, V., Baltezarević, R., & Milovanović, S. (2014). Between the lines and through the images. *Informatologija*, *47*(1), 29-35. https://hrcak.srce.hr/file/178309

Baltezarević, R., Baltezarević, B., Kwiatek, P., & Baltezarević, V. (2019). The impact of virtual communities on cultural identity. *Symposion, 6(*1), 7-22. https://doi.org/10.5840/symposion2019611

Baltezarević, R., & Baltezarević, I. (2024). Students' Attitudes on The Role of Artificial Intelligence (Ai) In Personalized Learning. *International Journal of Cognitive Research in Science, Engineering and Education (IJCRSEE), 12*(2), 387-397. http://dx.doi.org/10.23947/2334-8496-2024-12-2-387-397

Bareis, J., & Katzenbach, C. (2022). Talking AI into being: The narratives and imaginaries of national AI strategies and their performative politics. *Science, Technology, & Human Values, 47*(5), 855-881. https://doi.org/10.1177/01622439211030007

Barclay D. A. (2018). *Fake news, propaganda, and plain old lies: how to find trustworthy information in the digital age.* Lanham, MD: Rowman & Littlefield.

Battista, D., & Uva, G. (2023). Exploring the Legal Regulation of Social Media in Europe: A Review of Dynamics and Challenges—Current Trends and Future Developments. *Sustainability, 15*(5), 4144. https://doi.org/10.3390/su15054144

Bessi, A., & Ferrara, E. (2016). Social Bots Distort the 2016 US Presidential Election Online Discussion. *SSRN Electronic Journal.* https://ssrn.com/abstract=2982233

Bond, S. (2024, December). How AI deepfakes polluted elections in 2024. *NPR.* https://www.npr.org/2024/12/21/nx-s1-5220301/deepfakes-memes-artificial-intelligence-elections

Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences*, *114*(28), 7313–7318. http://dx.doi.org/10.1073/pnas.1618923114

Bradshaw, S., & Howard, P. N. (2018). *Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation.* Oxford University.

Brannon, W., Beeferman, D., Jiang, H., Heyward, A., & Roy, D. (2024). AudienceView: AI-assisted interpretation of audience feedback in journalism. *arXiv.* https://arxiv.org/abs/2407.12613

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A. & Agarwal, S. (2020). Language models are few-shot earners. *Advances in Neural Information Processing Systems, 33*, 1877-1901. https://arxiv.org/abs/2005.14165

BSI (2025). Digression: Social Bots and Chatbots. *Bundesamt für Sicherheit in der Informationstechnik.* https://www.bsi.bund.de/EN/Themen/Verbraucherinnen-und-Verbraucher/Informationen-und-Empfehlungen/Onlinekommunikation/Soziale-Netzwerke/Sichere-Verwendung/Exkurs-bots/social-bots.html

Burkhardt, J. M. (2017). History of Fake News. *Library Technology Reports, 53*(8), 5-9. https://journals.ala.org/index.php/ltr/article/view/6497/8631

Carr, N. G. (2011). *The shallows: What the internet is doing to our brains.* W. W. Norton & Company.

CBS News. (2019, May). Doctored Nancy Pelosi video highlights threat of "deepfake" tech. https://www.cbsnews.com/news/doctored-nancy-pelosi-video-highlights-threat-of-deepfake-tech-2019-05-25/

Chang, T. (2025, April). The AI Bot Epidemic: The Imperva 2025 Bad Bot Report. *Thales Group.* https://cpl.thalesgroup.com/blog/access-management/ai-bots-internet-traffic-imperva-2025-report

Chawla, R. (2019). Deepfakes: How a pervert shook the world. *International Journal of Advance Research and Development*, *4*(6), 4–8. https://www.semanticscholar.org/paper/Deepfakes-%3A-How-a-pervert-shook-the-world-Chawla/c3b3a6d-27dbbfed4df630b39fc0a8a6692b1828a

Cheguri, P. (2023). Deepfake Technology: Concerns Raised in the Advertising Industries. *Analytics Insight.* https://www.analyticsinsight.net/topic/deepfake-technology

Chester, J., & Montgomery, K. C. (2017). The role of digital marketing in political campaigns. *Internet Policy Review*, *6*(4), 1–20. http://dx.doi.org/10.14763/2017.4.773

Citron, D. K., & Pasquale, F. (2019). The scored society: Due process for automated predictions. *Washington Law Review, 89*(1), 1–33. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2376209

Coeckelbergh, M. (2022). *The Political Philosophy of AI: An Introduction.* John Wiley & Sons: New York, NY, USA.

Corney, D., Wilkinson, K., & Cann, R. (2024, June). The AI election: How Full Fact is leveraging new technology for UK general election fact checking. *Full Fact.* https://fullfact.org/blog/2024/jun/the-ai-election-how-full-fact-is-leveraging-new-technology-for-uk-general-election-fact-checking/

Crawford, K. (2021). *The atlas of AI: Power, politics, and the planetary costs of artificial intelligence.* Yale University Press. https://doi.org/10.12987/9780300252392

Crilley, R. (2018). International relations in the age of 'post-truth' politics. *International Affairs, 94*(2), 417-425. http://dx.doi.org/10.1093/ia/iiy038

Cruz, B. (2024, August). 2024 Deepfakes Guide and Statistics. *Security.* https://www.security.org/resources/deepfake-statistics/

Dudfield, A. (2025, August). How to stop AI chatbots going rogue. *Full Facts.* https://fullfact.org/technology/how-to-stop-ai-chatbots-going-rogue/

Ethayarajh, K., Xu, W., Muennighoff, N., Jurafsky, D., & Kiela, D. (2024). KTO: Model Alignment as Prospect Theoretic Optimization. *arXiv.* https://arxiv.org/abs/2402.01306

European Commission. (2021, April). Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain union legislative acts. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206

European Parliament. (2022, July). Digital Services: landmark rules adopted for a safer, open online environment. https://www.eu-

Baltezarević, R., Lović, V., & Baltezarević, I. (2025). Between Progress and Peril: The Role of Artificial Intelligence (AI) in Shaping Modern Political Communication, *International Journal of Cognitive Research in Science, Engineering and Education (IJCRSEE), 13*(3), 823-835.

roparl.europa.eu/news/en/press-room/20220701IPR34364/digital-services-act-eu-rules-to-make-digital-platforms-safer

Factiverse. (2025, May). How Factiverse Scans the Web to Tackle Misinformation at Scale. https://www.factiverse.ai/blog/how-factiverse-scans-the-web-to-tackle-misinformation-at-scale

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM, 59*(7), 96–104. https://dl.acm.org/doi/10.1145/2818717

Fetzer, J.H., & Fetzer, J.H. (1990). *What Is Artificial Intelligence?* Springer.

Fullfact. (2025). Find and Fight Bad Information. https://fullfact.ai/about/

Funk, A., Shahbaz, A., & Vesteinsson, K. (2023, November). The Repressive Power of Artificial Intelligence. *Freedom House.* https://freedomhouse.org/report/freedom-net/2023/repressive-power-artificial-intelligence

Gallo, M., Fenza, G., & Battista, D. (2022). Information Disorder: What about global security implications? *Rivista di Digital Politics, 2*(3), 523-538. https://doi.org/10.53227/106458

Gilbert, D. (2025, June). AI Chatbots Are Making LA Protest Disinformation Worse. *Wired.* https://www.wired.com/story/grok-chatgpt-ai-los-angeles-protest-disinformation

Gillham, J. (2025, September). Grover AI Content Detection Review. *Originality.ai* https://originality.ai/blog/grover-ai-content-detection-review

Gillespie, T. (2018). *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media.* Yale University Press.

Gorwa, R. (2019). The platform governance triangle: conceptualising the informal regulation of online content. *Internet Policy Review*, *8*(2). https://doi.org/10.14763/2019.2.1407

Guess, A.M., Lerner, M., Lyons, B., Montgomery, J.M., Nyhan, B., Reifler, J., & Sircar, N. (2020). A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. *Proceedings of the National Academy of Sciences of the United States of America, 117*(27), 15536-15545. https://doi.org/10.1073/pnas.1920498117

Hagerty, A., & Rubinov, I. (2019, July). Global AI ethics: A review of the social impacts and ethical implications of artificial intelligence. *arXiv.* https://arxiv.org/abs/1907.07892

Hetrick. C. (2024, July). How to spot AI fake news – and what policymakers can do to help. *USC Price School of Public Policy.* https://priceschool.usc.edu/news/ai-election-disinformation-biden-california-europe/

Hu, L., Wei, S., Zhao, Z., & Wu, B. (2022). Deep learning for fake news detection: A compre-hensive survey. *AI Open, 3,* 133-155, https://doi.org/10.1016/j.aiopen.2022.09.001

Islam, M.B.E., Haseeb, M., Batool, H., Ahtasham, N., & Muhammad, Z. (2024). AI Threats to Politics, Elections, and Democracy: A Blockchain-Based Deepfake Authenticity Verification Framework. *Blockchains 2024, 2*(4), 458–481. https://doi.org/10.3390/blockchains2040020

Jiang, J., Ren, X., & Ferrara, E. (2021). Social Media Polarization and Echo Chambers in the Context of COVID-19: Case Study. *JMIRx Med, 2*(3), e29570 https://doi.org/10.2196/29570

Kerly, R. (2020, August). How Deepfakes Are Changing Digital Marketing. *Loop Digital.* https://www.loop-digital.co.uk/the-rise-of-deepfake-technology/

Khare, Y. (2023, April). The Role of AI in Political Campaigns: Revolutionizing the Game. *Analytics Vidhya.* https://www.analyt-icsvidhya.com/blog/2023/04/the-role-of-ai-in-political-campaigns-revolutionizing-the-game/

Klinger, U., Kreiss, D., & Mutsvairo, B. (2023). Platforms, Power, and Politics: A Model for an Ever-changing Field. *Political Communication Report, 27*, 1-6. http://dx.doi.org/10.17169/refubium-39045

Konopliov, A. (2024, June). Key Statistics on Fake News & Misinformation in Media in 2024. *Redline Digital.* https://redline.digital/fake-news-statistics/

Korshunov, P., & Marcel, S. (2018, December). DeepFakes: a New Threat to Face Recognition? Assessment and Detection. *arXiv.* https://arxiv.org/abs/1812.08685

Kumar, A., & Garg, G. (2020). Systematic Literature Review on Context-Based Sentiment Analysis in Social Multimedia. *Multimedia Tools and Applications, 79(*21-22), 15349–15380. https://doi.org/10.1007/s11042-019-7346-5

Kumar, S. (2025, May). Why Governments Worldwide Are Enacting Stricter AI Deepfake Regulations in 2025. *Medium.* https://medium.com/@meisshaily/why-governments-worldwide-are-enacting-stricter-ai-deepfake-regulations-in-2025-32a61309366c

Landrin, S. (2024, May). India's general election is being impacted by deepfakes. *LeMonde.* https://www.lemonde.fr/en/pixels/article/2024/05/21/india-s-general-election-is-being-impacted-by-deepfakes_6672168_13.html

Lazaro Cabrera, L. (2024, May). AI Policy & Governance, European Policy, Free Expression. EU AI Act Brief – Pt. 3, Freedom of Expression. *Center for Democracy & Technology.* https://cdt.org/insights/eu-ai-act-brief-pt-3-freedom-of-expression/

Loth, A., Kappes, M., & Pahl, M.-O. (2024, April). Blessing or curse? A survey on the Impact of Generative AI on Fake News. *arXiv.* https://arxiv.org/abs/2404.03021

Lutkevich, B. & Hildreth, S. (2022, February). Social listening (social media listening). *TechTarget.* https://www.techtarget.com/searchcustomerexperience/definition/social-media-listening

Maheshwari, S. (2024, August). *Brands Love Influencers (Until Politics Get Involved).* The New York Times. https://www.nytimes.com/2024/08/12/business/media/influencers-politics-ai-analysis.html

Matheis, A. (2023). How can artificial intelligence be used in political communication? *Wegewerk.* https://www.wegewerk.com/en/blog/how-can-artificial-intelligence-be-used-in-political-communication/

Mermoud, A. (2017, August). The Power of Big Data and Psychographics in politics. *Swiss Intell.* https://swissintell.ch/the-power-of-big-data-and-psychographics-in-politics/

Murali, P., Hernandez, J., McDuff, D., Rowan, K., Suh, J., & Czerwinski, M. (2021, January). AffectiveSpotlight: Facilitating the communication of affective responses from audience members during online presentations. *arXiv.* https://arxiv.org/abs/2101.12284

NBC News (2023, December). Putin quizzed about AI and body doubles by his apparent deepfake. https://www.nbcnews.com/video/putin-quizzed-about-ai-and-body-doubles-by-his-apparent-deepfake-200210501620

Nida-Rumelin, J., & Weidenfeld, N. (2019). *Umanesimo digitale: un'etica per l'epoca dell'Intelligenza artificiale.* FrancoAngeli.

Noto, G. (2024, May). Scammers siphon $25M from engineering firm Arup via AI deepfake 'CFO'. *CFO Dive.* https://www.cfodive.com/news/scammers-siphon-25m-engineering-firm-arup-deepfake-cfo-ai/716501/

Nunziata, F. (2021). Il platform leader. *Rivista di Digital Politics, 1*(1), 127-146. https://www.rivisteweb.it/doi/10.53227/101176

Obot, O. U., Attai, K. F., Onwodi, G. O., James, I., & John, A. (2025). Sentiment analysis of electronic word of mouth (E-WoM) on e-learning. In M. Khosrow-Pour (Ed.), *Encyclopedia of Information Science and Technology (6th ed., ch. 57).* IGI Global. https://doi.org/10.4018/978-1-6684-7366-5.ch057

Pamment, J., Nothhaft, H., & Fjällhed, A. (2018). Countering information influence activities: A handbook for communicators. *Swedish Civil Contingencies Agency (MSB).* https://rib.msb.se/filer/pdf/28697.pdf

Parry, R. (2006, December). The GOP's $3 Bn Propaganda Organ. *The Baltimore Chronicle.* https://baltimorechronicle.com/

Pennycook, G., Cannon, T.D., & Rand, D.G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General, 147*(12), 1865-1880. https://doi.org/10.1037/xge0000465

Petrosyan, A. (2025, May). Potential influence of AI and deepfakes on upcoming elections 2024, by country. *Statista.* https://www.statista.com/statistics/1534957/global-potential-influence-ai-elections-by-country/

Political Communication. (2023). AI and Political Communication. *Political Communication Report*, 27(Spring). https://political-communication.org/article/ai-and-political-communication/

Populismstudies. (2018). Filter Bubbles. *ECPS.* https://www.populismstudies.org/Vocabulary/filter-bubbles/

Rajashekhargouda, P. (2022). Sentimental Analysis on Amazon Reviews Using Machine Learning. In Karuppusamy, P., García Márquez, F.P., & Nguyen, T.N., (Eds.). *Ubiquitous Intelligent Systems* (pp. 467–477). Springer Nature Singapore.

Roberts, S. T. (2019). *Behind the screen: Content moderation in the shadows of social media.* Yale University Press.

Roozenbeek, J., & van der Linden, S. (2020). Breaking Harmony Square: A game that "inoculates" against political misinformation. *Harvard Kennedy School (HKS) Misinformation Review, 1*(8). https://doi.org/10.37016/mr-2020-47

Samoili, S., López Cobo, M., Gómez, E., De Prato, G., Martínez-Plumed, F., & Delipetrev, B., (2020). AI watch. Defining artificial intelligence. Towards an operational definition and taxonomy of artificial intelligence (EUR 30117 EN). *Publications Office of the European Union.* https://doi.org/10.2760/382730

Schneider, B. (2017, June). How Vote Leave Used Data Science and A/B Testing to Achieve Brexit. *AB Tasty.* https://www.abtasty.com/blog/data-science-ab-testing-vote-brexit/

Seitz-Wald, A. (2024, February). A New Orleans magician says a Democratic operative paid him to make the fake Biden robocall. *NBC News.* https://www.nbcnews.com/politics/2024-election/biden-robocall-new-hampshire-strategist-rcna139760?_ga=2.181210351.976717714.1719011557-176973521.1719011550

Shen, T., Ruixian, L., Ju, B., & Zheng, L. (2018). 'Deep Fakes' Using Generative Adversarial Networks (GAN) (Report No. 16). *Noiselab, University of California, San Diego.* http://noiselab.ucsd.edu/ECE228_2018/Reports/Report16.pdf

Somers, M. (2020, July). Deepfakes, explained. *MIT Sloan School of Management.* https://mitsloan.mit.edu/ideas-made-to-matter/deepfakes-explained

Statista Research Department. (2024). U.S. adults worry about AI-generated political propaganda 2023. *Statista.* https://www.statista.com/statistics/1471069/us-adults-ai-generated-political-propaganda/

Surfshark. (2025). Deepfake statistics in early 2025: how frequently are famous people targeted? https://surfshark.com/research/study/deepfake-statistics

Sustainability Directory. (2025, May). How Effective Is Digital Literacy in Addressing Misinformation? https://sustainability-directory.com/question/how-effective-is-digital-literacy-in-addressing-misinformation/

Suzor, N. (2019). *Lawless: The secret rules that govern our digital lives.* Cambridge University Press. https://doi.org/10.1017/9781108666428

Tandoc, E. C., J., Lim, Z. W., & Ling, R. (2018). Defining "Fake News": A typology of scholarly definitions. *Digital Journalism, 6*(2), 137–153. https://doi.org/10.1080/21670811.2017.1360143

Thornhill, J. (2024, June). The danger of deepfakes is not what you think. *The Straits Times.* https://www.straitstimes.com/opinion/the-danger-of-deepfakes-is-not-what-you-think

Baltezarević, R., Lović, V., & Baltezarević, I. (2025). Between Progress and Peril: The Role of Artificial Intelligence (AI) in Shaping Modern Political Communication, *International Journal of Cognitive Research in Science, Engineering and Education (IJCRSEE), 13*(3), 823-835.

Verma, P., & De Vynck, G. (2024, January). AI is destabilizing 'the concept of truth itself' in 2024 election. *The Washington Post*. https://www.washingtonpost.com/technology/2024/01/22/ai-deepfake-elections-politicians

Virginia Tech. (2024, February). AI and the spread of fake news sites: Experts explain how to counteract them. *Virginia Tech News*. https://news.vt.edu/articles/2024/02/AI-generated-fake-news-experts.html

Viudes, F. J. (2023). Revolucionando la política: El papel omnipresente de la IA en la segmentación y el targeting de campañas modernas. *Más poder local, (53)*, 146-151. https://doi.org/10.56151/maspoderlocal.183

Vosoughi, S., Deb, R. & Aral, S. (2018). The spread of true and false news online. *Science, 359*(6380*)*,1146-1151. https://www.science.org/doi/10.1126/science.aap9559

Wakefield, J. (2019, November). Brittany Kaiser calls for Facebook political ad ban at Web Summit. *BBC News.* https://www.bbc.com/news/technology-50234144

Watson, A. (2024, January). Fake news in Europe - statistics & facts. *Statista.* https://www.statista.com/topics/5833/fake-news-in-europe/#topicOverview

Weforum (2024). Fake news undermines democracy, warns global survey. https://www.weforum.org/videos/influence-of-fake-news/

Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review, 9*(11), 40-53. https://doi.org/10.22215/timreview/1282

Yu, C. (2024). *How will AI steal our elections? (OSF Preprint, un7ev).* Center for Open Science. https://doi.org/10.31219/osf.io/un7ev

Zandt, F. (2024, March). How Dangerous are Deepfakes and Other AI-Powered Fraud? *Statista Daily Data.* https://www.statista.com/chart/31901/countries-per-region-with-biggest-increases-in-deepfake-specific-fraud-cases/

Zannettou, S., Sirivianos, M., Blackburn, J. & Kourtellis, N. (2019). The Web of False Information: Rumors, Fake News, Hoaxes, Clickbait, and Various Other Shenanigans. *Journal of Data and Information Quality, 1*(3), Article No. 10. https://doi.org/10.1145/3309699

Zellers, R., Holtzman, A., Rashkin, H., Bisk, Y., Farhadi, A., Roesner, F., & Choi, Y. (2019). Defending against neural fake news. *Advances in neural information processing systems, 32.* https://arxiv.org/abs/1905.12616