

Don't Panic (Yet): Assessing the Evidence and Discourse Around Generative AI and Elections

By Felix M. Simon and Sacha Altay
July 2025

Abstract

Prominent voices worry that generative artificial intelligence (GenAI) will negatively impact elections worldwide and trigger a misinformation apocalypse. A recurrent fear is that GenAI will make it easier to influence voters and facilitate the creation and dissemination of potent mis- and disinformation. We argue that despite the incredible capabilities of GenAI systems, their influence on election outcomes has been overestimated. Looking back at 2024, the predicted outsized effects of GenAI did not happen and were overshadowed by traditional sources of influence. We review current evidence on the impact of GenAI in the 2024 elections and identify several reasons why the impact of GenAI on elections has been overblown. These include the inherent challenges of mass persuasion, the complexity of media effects and people's interaction with technology, the difficulty of reaching target audiences, and the limited effectiveness of AI-driven microtargeting in political campaigns. Additionally, we argue that the socioeconomic, cultural, and personal factors that shape voting behavior outweigh the influence of AI-generated content. We further analyze the bifurcated discourse on GenAI's role in elections, framing it as part of the ongoing "cycle of

technology panics." While acknowledging AI's risks, such as amplifying social inequalities, we argue that focusing on AI distracts from more structural threats to elections and democracy, including voter disenfranchisement and attacks on election integrity. The paper calls for a recalibration of the narratives around AI and elections, proposing a nuanced approach that considers AI within broader sociopolitical contexts.

Keywords: Artificial Intelligence, Generative AI, Elections, Democracy, Misinformation, Political Campaigns, Voting Behavior

Introduction		
Background: Elections, Mass Persuasion, GenAl	8	
Definition of AI and GenAI	8	
Elections and voting behavior	9	
The limits of mass persuasion	11	
Common Arguments about GenAl and Risks to Elections	17	
Addressing the Main Claims about AI Risks around Elections	20	
1) AI will increase the quantity of information and misinformation around elections	20	
2) AI will increase the quality of election misinformation	24	
3) AI will improve the personalization of information and misinformation at scale	28	
4) New modes of information consumption	36	
5) Destabilization of reality	40	
6) Human-Al relationships	44	
Discussion	49	
1. What counts as AI-generated content?	49	
2. How do people actually use AI?	50	
3. How AI was used during elections	52	
Round and Round it Goes? Possible Reasons for the Skewed Discourse on AI and Elections	56	
The role of technological determinism	57	
Self-serving claims by industry, academia, and experts	59	
Competing levels of expertise	59	
Dynamics of the attention economy	60	

Political opportunism	61
The window of opportunity effect	62
Genuine concern and personal experience	62
Consequences of a Skewed AI Discourse	63
The focus on misuses of GenAI in elections can divert us from other harms	63
Narratives about the outsized effect of AI on elections could lead to suboptimal policy responses	65
The alarmist discourse on the effect of AI on elections could reduce trust and satisfaction with democracy	65
Conclusion	67
Research Agenda	70
Acknowledgments	74
References	75
About the Authors	92
About the Knight First Amendment Institute	93

Introduction

The increasing public availability of generative artificial intelligence (GenAI) systems, such as OpenAI's ChatGPT, Google's Gemini, and a slew of others has led to a resurgence of concerns about the impact of AI and GenAI in public discourse. Leading voices from politics, business, and the media twice listed "adverse outcomes of AI technologies" as having a potentially severe impact in the next two years (together with "mis- and disinformation") in the World Economic Forum's Global Risks Reports (2024, 2025). The public is worried as well. A recent survey of eight countries, including Brazil, Japan, the U.K., and the U.S. found that 84 percent of people were concerned about the use of AI to create fake content (Ejaz et al., 2024). Meanwhile, a large survey of AI researchers found that 86 percent were significantly or extremely concerned about AI and the spread of false information, and 79 percent about manipulation of large-scale public opinion trends (Grace et al., 2024). The main worry present in all these contexts is that AI will make it easier to create and target potent mis- and disinformation and propaganda and manipulate voters more effectively. The integration of foundation models, particularly AI chatbots, into various digital media and their growing use for online searches, interaction with information and news, and use as personal assistants is also a growing concern, due to the potential knock-on effects on people's informedness about politics and their political behavior.

A recurrent theme is the impact of AI on national elections. Initial predictions warned that GenAI would propel the world toward a "tech-enabled Armageddon" (Scott, 2023), where "elections get screwed up" (Verma & Zakrzewski, 2024), and that "anybody who's not worried [was] not paying attention" (Aspen Digital, 2024). We critically examine these claims against the backdrop of the 2023-2024 global election cycle, during which nearly half of the world's population had the opportunity to participate in elections, including in high-stakes contests in countries such as the U.S. and Brazil.

We make three contributions: *First*, we argue that despite widespread predictions of AI-driven electoral manipulation through, for example, deepfakes, as well as AI-informed targeted advertising and misinformation campaigns, the influence of AI on national elections was largely overshadowed by other, much more important factors, such as politicians' willingness to misinform, lie, and break other norms. We identify several key factors contributing to this discrepancy between alarmist predictions and observed outcomes:

- 1. The inherent challenges of mass persuasion, regardless of the tools employed
- 2. The difficulty of reaching target audiences in an oversaturated information landscape and high-choice media environments
- Emerging evidence regarding the limited effectiveness and use of AI-driven microtargeting in political campaigns
- 4. The complex interplay of socioeconomic, cultural, and personal factors that shape voting behavior, which often outweighs the influence of AI-generated content
- 5. The ways in which people consume information and decide who to trust and who to listen to

We argue that these factors explain why the worst predictions about the role of GenAI in recent national elections did not come to pass and should make us skeptical about claims that GenAI will upend elections in the years to come. Most of this is settled knowledge and based on long scholarly traditions in media effects studies and political science. For these reasons, we were skeptical of the impact of AI on elections even before the broader realization that the AI apocalypse did not occur (Simon et al., 2023).

Second, we provide an overview of the possible reasons for the skewed discourse on AI and elections. We diagnose the alarmist discourse surrounding AI and elections as a version of what Orben (2020) has termed the "Sisyphean cycle of technology panics"—the repeating, ultimately unproductive pattern of public and institutional overreaction to new

technologies—with several concurrent push and pull factors within political, technological, regulatory, media, and academic communities fostering this narrative.

Third, we interrogate the possible consequences of a skewed understanding of AI's impact on national elections and align it with—at times overlooked—other risks to elections and democracy, both from advances in AI and other, more structural and long-standing factors. While acknowledging the potential risks associated with AI in electoral contexts, such as the amplification of existing social inequalities and the erosion of trust in democratic processes, we argue that the disproportionate focus on GenAI may distract from other, more pressing threats to democracy. These include, but are not limited to, forms of voter disenfranchisement; unequal electoral competition, including in the access to digital tools; intimidation of election officials; attacks on journalists and politicians; and various forms of state oppression.

We argue that the current narrative surrounding GenAI's impact on elections, especially in democratic systems, requires recalibration. We propose a more nuanced approach to understanding the role of AI in electoral processes, one that considers the broader sociopolitical context and avoids both outright technological determinism and extreme social construction of technology arguments. We try to acknowledge the effects of AI within limits, while recognizing the social and political conditions that constrain it.

As such, this paper makes the following contributions to ongoing debates about the impact of AI on democracy:

- Integrating AI into existing theories and within the empirical literature in political communication and political science;
- 2. Staking out a research agenda for the future study of this topic;
- 3. Offering insights for policymakers, journalists, and researchers concerned with preserving the integrity of electoral processes in the "age of GenAI."

Background: Elections, Mass Persuasion, GenAl

Before focusing on the role of GenAI in elections, we provide a definition and summary of key developments in GenAI, before turning to what elections are and what shapes voting behavior. We then summarize key literature on mass persuasion—all of which will be integral to our later arguments.

Definition of Al and GenAl

AI broadly refers to the development and application of computer systems or machines capable of performing tasks that typically require human intelligence, including learning, reasoning, natural language processing, problem-solving, and decision-making (Mitchell, 2019; Russell & Norvig 2009). The OECD offers a comprehensive definition, describing an AI system as a "machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments" (OECD AI Principles, 2024). These systems vary in their levels of autonomy and adaptability after deployment. Recent advancements in AI have led to the development of "general-purpose AI models", designed to be adaptable to a wide range of downstream tasks (Bengio et al., 2025), for example through additional fine-tuning. Foundation models are usually large-scale neural networks trained on diverse data sets and capable of performing a wide range of tasks across multiple modalities, including text, code, visuals, and audio. Large language models (LLMs) in particular are a subset of foundation models specialized in text processing and generation (House of Lords Communications and Digital Committee: Large Language Models and Generative AI, 2024).

The term 'generative AI' has seen a substantive rise in coverage since 2022

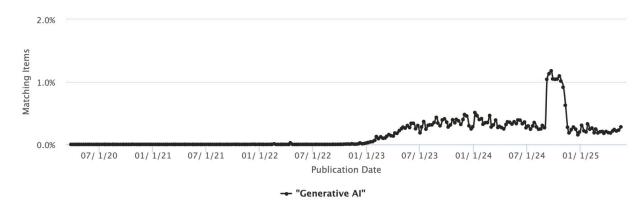


Figure 1. This chart illustrates the percentage of total coverage of stories (weekly) matching the term "Generative AI," drawn from approximately 1,500 English-speaking media sources. Source: Global English Language Sources database, provided by MediaCloud, spanning the period from April 1, 2020, to May 21, 2025.

The term 'generative AI' (or 'GenAI') specifically became prominent following the public release of the AI system ChatGPT—a chatbot developed by U.S. firm OpenAI—on November 30, 2022. The term loosely refers to AI systems capable of rapidly creating new data, including content across various modalities and formats that is often perceived as indistinguishable from human-generated content (e.g., in the case of text) or content generated with other analog or digital means (e.g., audio or video), depending on the instructions provided. The launch of ChatGPT at the end of 2022 led to a wave of similar releases from major technology companies, including Google and Meta and a range of smaller firms, who introduced their own chatbots and AI models in quick succession. This trend, and the fact that these systems can be accessed and used through natural language interfaces, has significantly expanded the accessibility and usefulness of AI systems for general users.

Elections and voting behavior

We should clarify that this paper focuses on the effects of GenAI in broadly democratic systems. While some of the arguments may be applicable to electoral or closed autocracies,

these are not the focus of this article. There is ongoing debate in political science about the exact nature of democracy, a discussion that also falls outside the scope of this paper. Instead, we adopt a minimalist definition that defines democracy as a system of government in which power is vested in the people and rulers are elected through competitive elections. Elections, in this context, are a mechanism by which political conflict in society is channeled into real power over society within an institutional framework for a given amount of time (see also Jungherr, 2023).

People do not approach elections as blank slates, waiting to be persuaded by the latest piece of information or political arguments. Instead, voting behaviors are shaped by a complex nexus of factors, usually a mixture of long-standing predispositions and short-term contextual factors (Campbell et al., 1960; Zaller, 1992). Classic models suggest that partisanship, often formed early in life, serves as a primary filter through which individuals process political information and cast their ballots (Green et al., 2002). Identity-based considerations—such as social group attachments and ideological orientations—can interact with media exposure to shape voters' impressions of and feelings toward candidates and policy issues (Huddy et al., 2023, Tenenboim-Weinblatt et al., 2022).

Political behavior, in turn, reflects the interplay of these attitudes with external mobilization efforts, social networks, and media frames.¹ For example, socioeconomic status not only correlates strongly with political engagement but also shapes individuals' sense of political efficacy and their capacity to mobilize effectively (Brady et al., 1995; Oser et al., 2022). Get-out-the-vote campaigns, peer group discussions, and opinion leadership can all prompt individuals to (dis)engage in political activities, such as attending rallies or casting a ballot (Hansen & Pedersen, 2014). In modern contexts, social media platforms

Media frames are the interpretive structures that news producers use to select, emphasize, and organize elements of reality so audiences understand an issue in terms of specific problem definitions, causal explanations, moral judgments, and potential remedies.

intensify this dynamic: They amplify message exposure, facilitate political discussion, and are widely used by political actors to stay informed, disseminate information, and engage voters, while voters actively or passively consume information or participate in political actions on these same platforms. However, not all groups experience these effects equally; systemic factors like socioeconomic status and educational resources can either enable or constrain political participation (Leighley & Nagler, 2013), and not all people have access to or are engaged with digital media where politics and political discourse play out.

The limits of mass persuasion

Political campaigns, along with commercial advertising, represent the most ambitious efforts at mass persuasion. This persuasion can have several aims: increasing or decreasing political participation (e.g., in the form of turnout in general, or for or against a particular side), increasing or decreasing political activism (e.g., raising money or political support during rallies), and shaping voting decisions. Elections are high-stakes events where candidates, parties, and interest groups pour—sometimes enormous—resources into persuading voters. For example, during the 2024 U.S. presidential election, \$1.35 billion was spent on online campaign ads on Google and Meta (Brennan Center for Justice, 2024), while more than \$15 billion was spent in the whole election cycle (Federal Election Commission, 2025). Despite these massive investments, numerous high-quality studies have shown that the persuasive effects of political advertising are limited in the U.S. (Allcott et al., 2025; Coppock et al., 2020, 2022; Haenschen, 2023) and elsewhere (Hager, 2019). For instance, Kalla and Broockman (2017) write that "the best estimate of the effects of campaign contact and advertising on Americans' candidates choices in general elections is zero. First, a systematic meta-analysis of 40 field experiments estimates an average effect of zero in general elections. Second, we present nine original field experiments that increase the statistical evidence in the literature about the persuasive effects of personal contact tenfold. These

experiments' average effect is also zero" (p. 1). Increasing statistical power to detect ever smaller effects, an eight-month political advertising campaign on social media delivered to 2 million persuadable voters found that "persuasion campaigns can indeed cause small differential turnout effects—much smaller than pundits and media commentators often assume, but our field experimental study is large enough to show that these effects are distinct from zero" (Aggarwal et al., 2023, p. 335). A large-scale experiment conducted on Facebook and Instagram, which removed political ads for six weeks prior to the 2020 U.S. presidential election, found null effects on political knowledge, polarization, perceived legitimacy of the election, political participation, candidate favorability, and turnout (Allcott et al., 2025).

Research on advertising more broadly has shown that ads have small effects on consumers, far smaller than commonly assumed (Della Vigna & Gentzkow, 2010; Hwang, 2020, Shapiro et al., 2021). For instance, a meta-analysis of 751 short-term and 402 long-term direct-to-consumer brand advertising elasticities showed that "the mean short-term advertising elasticity across all observations (1940-2004) is .12, the median elasticity is .05, and elasticity is declining over time. The finding that advertising elasticity is 'small' may upset many practitioners, especially those in the agency business" (Sethuraman et al., 2018, p. 469). "Elasticity" here quantifies how much 'bang' the advertiser gets for their 'buck,' so to speak, and this bang is surprisingly small. Raising ad expenditure by one percent typically boosted sales by only about one-tenth of a percent, and this was also declining. In other words: Each additional dollar of advertising is buying even *less* incremental consumer response than it once did. More generally, most attempts at mass persuasion fail to influence people's behaviors in meaningful ways or have negligible effects when they are not well aligned with their incentives (the things that motivate them), preferences (what they like or dislike), or values (their principles or standards of behavior; Mercier 2020). For instance, public health campaigns encouraging people to eat healthier or quit smoking have little

impact compared to making healthy options more affordable or convenient, or increasing the price of cigarettes (Arno & Thomas, 2016; Bader et al., 2011).

The fields of communication and political science have long moved away from models of mass communication which assumed strong, direct effects on public opinion (such as the "hypodermic needle model"). Instead, contemporary research favors more nuanced models of influence—such as agenda-setting and framing—that highlight smaller, indirect effects (Bryant & Oliver, 2009). Research on the limited influence of mass persuasion clashes with the widespread assumption that people are gullible and easily persuaded. It has been well documented that we perceive others to be much more susceptible to negative media effects than ourselves (such as propaganda or misinformation, but also pornographic content; Davison, 1983). This "third -person effect" also plays a central role in fears about the effect of misinformation (Altay & Acerbi, 2024) and may explain why fears about the effect of GenAI on misinformation and persuasion during elections are so widespread. Rather than being overly gullible, people are often quite resistant to changing their views. When people come across new information, they tend to update their beliefs in the direction of the new information, that is, they adjust their views slightly to reflect what they have just learned. However, this updating is often minimal and rarely leads to lasting shifts in attitudes or behaviors (Coppock, 2023), especially in political contexts. For instance, numerous studies have demonstrated that while fact-checking (in itself an attempt at persuading people to hold correct beliefs) reliably reduces misperceptions, it has a negligible influence on political attitudes (e.g., feelings toward a candidate) and behaviors (e.g., voting intentions; Nyan et al., 2019; Porter & Wood, 2024). More broadly, research on social learning shows that people consistently favor their own intuitions, experiences, and beliefs, over information communicated by others – even when they have no reason to trust themselves more than others (Bailey et al., 2023). Rather than being overly influenced by communicated information, we frequently waste it (Morin et al., 2021). Finally, a meta-analysis on news discernment across

40 countries and more than 194,000 participants has shown that people are not gullible: on average people are (very) good at spotting false news (Pfänder & Altay, 2025). Yet, while people can tell true from false news, they tend to be excessively skeptical of true news and to err on the side of skepticism rather than credulity (Pfänder & Altay, 2025)—all of which makes persuasion harder.

A skeptical reader may wonder: What about the effect of repeated exposure to persuasive content over long periods of time? In line with this argument, the literature in experimental psychology has shown that people are more likely to believe repeated claims (Pillai, Fazio, & Effron, 2023). However, there are reasons to doubt that this "illusory-truth effect" translates well outside of experimental settings. For instance, while repeated exposure to false statements professed by Donald Trump increased belief in the statements among Republicans, it decreased belief in the statement among Democrats (Pillai, Kim, & Fazio., 2023). Moreover, the illusory truth effect disappears if the source is not trusted (Orchinik et al., 2024) or if the information environment is perceived as being low-quality (Orchinik et al., 2025). While it is difficult to estimate the causal effect of repeated exposure to persuasive content over time, a large study by Aggarwal et al. (2023), which exposed 1,711,547 persuadable voters to an average of 754 political ads over eight months, nonetheless found minimal impact on voters.

Skeptical readers may also argue that elites, such as politicians, play a dominant role in shaping public opinion. While it is true that they can act as agenda setters, influencing which issues receive attention and how they are framed, this influence is not unidirectional but reciprocal (Gilardi et al., 2022). Elite cues help individuals—especially those with little political knowledge—navigate complex issues by providing simplified narratives and partisan heuristics (Bullock, 2020). These cues can influence policy preferences, but their influence is limited. People stop following party cues if they go too strongly against their self-interest (Slothuus & Bisgaard, 2021) or deviate too much from their own political stances

(Mummolo et al., 2019). Moreover, party cues mostly cause people to behave as if they were better informed (i.e., a party cue is often a valid information shortcut; Tappin & McKay, 2021), they do not dominate the policy information people already hold, such that people do not blindly follow party cues, but instead rely on the substance of the policy—even when exposed to party cues (Bullock, 2011, 2020), and finally party cues do not reduce receptivity to arguments and evidence (Tappin et al., 2023). In general, individuals are not passively exposed to information; they actively seek out and engage with content that aligns with their prior beliefs, which limits the media and elites' ability to profoundly and durably change minds (Arceneaux & Johnson, 2013).

In this section, we have mostly framed persuasion in negative terms, but persuasion is not inherently negative and is an essential element of democracy. It is through mass persuasion that societies build consensus, resolve disputes, and make collective decisions without resorting to force. In many ways, and in specific circumstances, mass persuasion can work. When Russia attacked Ukraine in 2022, or when Queen Elizabeth II died, people quickly updated their beliefs, and acknowledged that the Queen was no longer alive and Ukraine no longer at peace. Similarly, people accept unintuitive scientific knowledge about the formation of continents or the laws of physics. Most of this happens simply because we trust journalists and scientists on these matters and that their positions converge—this trust is largely built through the demonstration of competence and reliability of these experts (Mercier, 2020). Even self-described populists—often assumed to be particularly resistant to expertise—seem to engage similarly. Peresman et al. (2025) found that while populists are less willing to accept expert advice, both populists and non-populists are equally responsive to strong arguments and expert source characteristics—i.e., they are more likely to accept advice when it is supported by strong arguments. GenAI is unlikely to benefit from this trust afforded to experts yet—for instance, only 27 percent of people trust GenAI for news and information about politics, while only seven percent actually use GenAI for news

and information about politics (Ejaz et al., 2024). In the future, as AI use grows and LLMs become more accurate, they are likely to benefit from this trust—although their persuasiveness will remain dependent on the external sources they rely on and the perceived agreement with sources that people find reliable.

Mass persuasion can occur and be effective not only when the source is highly trusted but also when people are exposed to strong arguments. For example, most people radically shift their views when shown a clear demonstration of the correct solution to a logical or factual problem (Mercier & Claidière, 2022). LLMs can generate and summarize effective arguments. Growing evidence suggests that discussions with LLMs can significantly change people's opinions on a wide range of topics: They can reduce belief in conspiracy theories (Costello et al., 2024), reduce concerns about HPV vaccination (Xu et al., 2025), increase pro-climate attitudes (Czarnek et al., 2025), or even reduce prejudice toward undocumented immigrants (Costello et al., 2025). While it is clear that LLMs are persuasive, they are not necessarily more effective than the most effective existing messages (Chen et al., 2025; Hackenburg & Margetts, 2024a; Sehgal et al., 2025). In addition, the cost of exposing audiences to LLMs' persuasive messages is currently greater than commonly used persuasive methods. As Chen et al. (2025) note, "it is currently much easier to scale traditional campaign persuasion methods than LLM-based persuasion" (p. 1) and "while AI-based persuasion can match human performance on a per-person basis under ideal conditions of forced exposure, its real-world deployment is currently constrained by exposure costs and audience sizes" (p. 13).

If LLMs can be used to persuade people of socially desirable viewpoints, couldn't bad actors also use them to convince people of harmful or undesirable viewpoints? The answer is 'yes,' but early evidence suggests that LLMs are persuasive because they provide factual information and targeted counterarguments (Costello et al., 2025). For instance, while the LLM in the study about conspiracy theories was extremely effective at reducing belief in

false conspiracy theories, it did not reduce belief in true conspiracies (such as the MK Ultra covert CIA program; Costello et al., 2024). Thus, on average, LLMs may be more persuasive when advocating for positions that are solidly backed up by evidence. It remains an open question to what extent the persuasiveness of LLMs' arguments correlates with the factual accuracy of the conclusions they advocate (Jones & Bergen, 2024). There are reasons to think that, in general, people recognize good arguments and are more receptive to factual information (Mercier & Sperber, 2017); however, LLMs are able to cherry-pick facts to advocate for dubious positions. That being said, the evidence discussed here is based on experiments where participants were paid to interact with LLMs. It remains unclear whether people would engage with LLMs in the same way outside of these experimental settings. In addition, people will find reasons to discount information they do not agree with and limit their exposure to such information. For instance, Wikipedia is widely distrusted among conspiracy theorists circles because they perceive it as being politically biased—for better or worse, the same will happen with LLMs.

In sum, people are not gullible, and mass persuasion is difficult, especially when messages do not align with people's existing incentives, preferences, or values. People are not passively exposed to information and selectively expose themselves to sources they agree with and trust.

Common Arguments about GenAl and Risks to Elections

According to various voices, including some leading AI researchers, GenAI will upend elections and pose a major threat to democracy. These arguments, we suggest, can be divided into six broad categories (Table 1). We do not claim that this list is exhaustive. Instead, it is a categorization of the most common arguments with concerns ranging from the content side (what voters see and hear) to the infrastructure side (the integrity of the voting systems) and the broader societal impact of eroding trust in democratic processes.

Table 1. Six arguments about the assumed negative impact of GenAI on elections.

Argument	Explanation of claim	Presumed effect	Source
1. Increased quantity	Due to its technical	Increased reach of	Bell (2023), Fried
of information and	capabilities and ease	misinformation	(2023), Hsu
misinformation	of use, GenAl can		& Thompson (2023),
	be used to create	Increased reach for	Marcus
	information at scale	political actors	(2023), Pasternack
	with great ease.		(2023),
		Crowding out of other	Ordonez et al. (2023),
		information	Tucker
			(2023), Zagni &
			Canetta
			(2023), Safiullah &
			Parveen (2022)
2. Increased quality	Due to its technical	Increased persuasion	Fried (2023), Gold
of information and	capabilities and	of voters (all	& Fischer (2023),
misinformation	ease of use, GenAl	information), increased	Ordonez et al. (2023),
	can be used to	susceptibility of voters	Pasternack (2023),
	create information	(misinformation)	Shah & Bender (2023),
	or misinformation		Zagni & Canetta
	perceived to be of		(2023), Safiullah &
	high quality at low		Parveen (2022)
	cost.		
3. Increased	Due to its technical	Increased persuasion of	Benson (2023),
personalization of	capabilities and ease	voters	Fried (2023), Hsu &
information and	of use, GenAl can		Thompson (2023),
misinformation at	be used to create		Pasternack (2023),
scale	high-quality (mis)		Safiullah & Parveen
	information at scale		(2022)
	personalized to a		
	user's tastes and		
	preferences.		

4. New modes	The integration of	Higher likelihood of	Angwin et al. (2024),
of information	GenAl into existing	voters being misinformed	Marinov (2024),
consumption	digital infrastructures	through GenAI	Simon, Fletcher,
	and the increasing		& Nielsen (2024),
	use of GenAl for	Provision of lower-quality	Rahman-Jones (2025),
	information seeking	or biased information	Safiullah & Parveen
	leads to changing	about elections	(2022), Jaźwińska &
	consumption patterns		Chandrasekar (2025)
	around election	Crowding out and long-	
	information.	term undermining of	
		authoritative sources of	
		election information	
5. Destabilization of	The realism of GenAl	Fostering undue	Goldstein & Lohn
reality	content creates	skepticism toward	(2024), Dowskin
	uncertainty regarding	accurate information	(2024), Carpenter
	what is real.		(2024), West & Lo
		Decline in public	(2024)
		trust and institutional	
		legitimacy	
		Weaponization to deny	
		inconvenient truths	
		("liar's dividend")	
6. Human-Al	Users form deeper,	Increased persuasion of	Knight (2016),
relationships	more persistent	voters	France24 (2025), Kirk
	relationships with		et al. (2025)
	personalized and	Increased misinformation	
	agentic AI systems.	of voters	
		Manipulation of voters	

It should be noted that these categories are not mutually exclusive and overlap in some cases. We use "information" here as a general term that puts a lesser emphasis on the exact

mode of delivery (e.g., advertisements in various digital media versus campaign messages published or uttered by candidates). In the section below, we discuss these claims in turn, arguing that current concerns about the effects of GenAI on elections are overblown in light of the available evidence and theoretical considerations. Afterward, we turn to a broader discussion to address common objections and examine the available evidence on how GenAI has been used in recent election cycles.

Addressing the Main Claims about AI Risks around Elections

In the following section, we will address each of the claims outlined above, based on recent evidence and the wealth of preexisting literature on technological change. One challenge in making claims about the role of new technologies in politics is the time lag between their emergence, their initial effects, and the availability of empirical research assessing these effects (see Orben, 2020, p. 1150f). While various empirical studies have examined the effects of GenAI in a political context at the time of writing, further research is needed. However, we argue that adequate expectations for the effects of GenAI can be formed from this new empirical material and the extensive existing literature on the role of digital technologies in elections (see also Dommett & Power, 2024).

1. AI will increase the quantity of information and misinformation around elections

The increase in the quantity of information, and especially misinformation, caused by GenAI uses could have various consequences, from polluting the information environment and crowding out quality information to swaying voters (see Table 1). Below, we dissect this argument.

The first premise is that GenAI will enhance the production of misinformation more than

that of reliable information. If GenAI primarily supports the creation of trustworthy content, its overall impact on the information ecosystem would be positive. However, it is difficult to quantify AI's role in content creation. The ways in which AI is used to support the production of reliable information (Simon, 2025) may be more subtle than AI uses to produce false information, and thus more difficult to detect. For the sake of argument, we assume that AI will be used exclusively to produce misinformation.

The second premise is that AI-generated content not only exists but also reaches people and captures their attention. This is the main bottleneck: For AI misinformation (and for information in general) to have an effect, it must be seen. On its own, (mis)information has no causal power. Yet, attention is a scarce resource, and the amount of information people can meaningfully engage with is finite, because time and attention are finite (Jungherr & Schroeder, 2021; Taylor, 2014), and any piece of politically relevant information has to compete with other types of information, such as entertainment. During elections, voters are already overwhelmed with messages and ads, making any additional content—AI-generated or not—another drop in the ocean. In low-information environments or data voids (boyd & Golebiewski, 2018), where fewer messages circulate, AI-generated content could have a stronger impact. But even in such cases, it is unclear why GenAI content would outperform authentic content and other non-AI-generated content. If anything, the proliferation of AI content may increase the value and demand for authentic content (see the Discussion). The third premise is that, after reaching its audience, AI content will be persuasive in some way. While strong direct effects on voter preferences are unlikely (see the section on mass persuasion and the determinants of voting), weak and indirect effects are conceivable. Elections are not only about votes but also about the quality of political debate and media coverage. Without swaying voters, AI could flood the information environment in ways that degrade public discourse and democratic processes. We do not find the flooding argument particularly convincing. Since most people rely on a small number of trusted sources for

news and politics (Newman et al., 2024; 2025), misleading AI content from less credible sources would likely have limited influence, if it gets seen at all. If mainstream media and trusted news influencers do not misuse AI, why would AI-generated content from untrusted actors cause confusion? Flooding is concerning when there is significant uncertainty about who to trust and when individuals lack control over their information exposure. Yet, in practice, international survey research shows that people still often turn to mainstream media despite low levels of trust and an expanding pool of content creators (Fletcher et al., 2025; Newman et al., 2024; 2025; Strömbäck et al., 2020). In Western democracies, mainstream news outlets have so far shown restraint in their use of GenAI. While some media organizations are using AI to assist in news production and distribution (Simon, 2024, 2025) these uses have generally been transparent and responsible. There is little evidence to suggest that mainstream news organizations are using AI to create misleading content or fake news. In fact, many news organizations are taking steps to ensure that AI-generated content is clearly labeled and that editorial oversight remains in place (Becker et al., 2025). This responsible approach stands in stark contrast to fears that AI could dominate political news coverage and create mass disinformation. Instead, news organizations are leveraging AI tools to enhance journalistic processes, such as fact-checking or summarizing data-heavy reports, rather than misleading their audiences.

The rise of news influencers, too, does not necessarily indicate a breakdown of the information ecosystem, although it presents a shift. In France, for example, HugoDécrypte, the country's largest news influencer, has grown into a respected media entity with mainstream, high-quality coverage. Trusted sources, whether they are news influencers or news outlets, have strong reputational incentives to appear credible, as their audience's trust—and their reach—largely depends on it (Altay et al., 2022a). Using AI to mislead—and being exposed by competitors—would be a death sentence for most news sources. These imperfect but powerful reputational incentives largely explain why we generally try to avoid

spreading falsehoods despite the ease of writing false text or making false claims (Sperber et al., 2010).

In general, fears about AI-driven increases in misinformation are missing the mark because they focus too heavily on the *supply* of information and overlook the role of *demand*. People consume and share (mis)information that aligns with their worldviews and seek out sources that cater to their perspectives (Arceneaux & Johnson, 2013). Motivated reasoning,² group identities, and societal conflict have been shown to increase receptivity to misinformation (Mazepus et al., 2023). For example, those with unfavorable views about vaccines are much more likely to visit vaccine-skeptical websites (Guess et al., 2020). Sharing misinformation is often also a political tool, with especially radical-right parties resorting to it "to draw political benefit" (Törnberg & Chueri, 2025). And some people consume and share false information as a result of social frustrations, seeking to disrupt an "established order that fails to accord them the respect that they feel they personally deserve" and in hopes of gaining status in the process (Petersen et al., 2023). Moreover, people do not even need to believe misinformation deeply to consume and share it, with some people sharing news of questionable accuracy because it has qualities that compensate for its potential inaccuracy, such as being interesting-if-true (Altay et al., 2021). In addition, the people most susceptible to misinformation are not passively exposed to it online; instead, they actively search for it (Motta et al., 2023; Robertson et al., 2023). Misinformation consumers are not unique because of their special access to false content (a difference in supply) but because of their propensity to seek it out (a difference in demand). The fact that demand drives misinformation consumption and sharing is perhaps the most important lesson from the misinformation literature. As Budak et al. write: "In our review of behavioural science research on online misinformation, we document a pattern of low exposure to false and inflammatory

² Motivated reasoning is the willingness to accept information that supports preexisting beliefs and, conversely, to reject and counterargue information that challenges those same beliefs.

content that is concentrated among a narrow fringe with strong motivations to seek out such information" (2024, p. 1).

To wrap up, the presence of more misinformation due to GenAI does not necessarily mean that people will consume more of it. For an increase in supply to translate into additional effects, there must be an unmet demand or a limited supply. Yet, the internet already contains plenty of low-quality content—much of which goes unnoticed (Budak et al., 2024). The barriers to creating and accessing misinformation are already extremely low, and we see no good reason to assume that people will show a higher demand for AI-generated misinformation over existing forms of misinformation. Throughout history, humans have shown a remarkable ability to make up false stories, from urban legends to conspiracy theories. Misinformation about elections is easy to create. All it requires is taking an image out of context, slowing down video footage, or simply saying plainly false things. In these conditions, GenAI content has very little room to operate. Moreover, the demand for misinformation is easy to meet: Misinformation sells as long as it supports the right narrative and resonates with people's identities, values, and experiences (more on this below).

2. AI will increase the quality of election misinformation

The quality of AI-generated (mis)information has sparked concerns about its potential to deceive people and erode trust in the information environment. By quality, we mean that GenAI enables the creation of text, imagery, audio, and video with such lifelike fidelity that observers cannot reliably distinguish these synthetic creations from material produced through conventional human activity—whether written by an author, photographed with a camera, or recorded with a microphone. Below, we examine this position and argue that while information quality is crucial in many contexts—such as determining guilt in a criminal case—it plays a much smaller role in the acceptance and spread of misinformation,

including in the context of elections.³ The core premise of the argument is that the danger to elections stems from GenAI making election misinformation more successful and impactful by improving its quality—and that AI could facilitate this by lowering the costs of producing high-quality misinformation.

Let us look at this argument step by step. First, it is undoubtedly true that GenAI models can now produce high-quality misinformation. Plausible but false text, audio, and visual material are all within the purview of recent, widely available and usable AI systems. Will these AI systems be used to produce more high-quality misinformation than reliable information? Possibly, not least due to different reputational incentives. News organizations and public-interest organizations depend on audience trust and brand reputation, so overt reliance on AI can backfire due to audience skepticism around AI (Nielsen & Fletcher, 2024; Toff & Simon, 2024) and worries about damage to their trust and reputation from errors AI systems still make. As a result, many outlets are still cautious in their AI use (Borchardt, 2024; Radcliffe, 2025), which curbs its potential to be used too widely for fully AI-generated high-quality, true information. By contrast, misinformation producers face no comparable reputational constraint: faster, cheaper, and less labor-intensive production of high-quality material is an advantage for actors with little to lose from being caught fabricating content. Moreover, professional journalism still enjoys stronger financial and institutional support than most misinformation operations, so the marginal value of additional cost reductions in producing high-quality content is lower for newsrooms than for bad-faith actors. For these reasons, we assume that reductions in the cost and turnaround time of high-quality

The case that an increase in the quality of information with the help of GenAI that is true and not meant to mislead poses no problem to elections is quickly made: It would benefit voters. Higher quality, true information provides a better basis to inform voting decisions. Here, the only quibble one might have is that GenAI could benefit some parties or candidates over others in the short to medium term, with those with greater resources better able to use AI to produce, for example, higher-quality, true campaigning material. However, the nature of political contest has always been beset by a struggle for a competitive advantage, and political actors have traditionally adapted to forms of innovation established by others (Jungherr et al., 2020).

outputs will confer proportionally greater benefits on misinformation producers than on reliable publishers—though GenAI can and is being harnessed to support responsible journalism as well (Simon, 2025).

However, this does not need to make a significant difference in the context of elections, for various reasons. First, while GenAI undoubtedly enables the creation of more sophisticated false content and might benefit its producers more, it is not clear that higher-quality misinformation would actually be more successful in persuading or misleading people. As we have already discussed, other factors—such as a demand for misinformation, as well as ideological alignment, emotional appeal, resonance with personal experiences, and the source—matter in determining who and why people accept and share misinformation. In other words: It is not just content quality that determines the spread and influence of misinformation. These factors will not be simply overwritten by an increase in content quality, and higher quality does not automatically lead to increased demand for misinformation. Second, misinformation producers already have numerous tools in their arsenals to enhance content quality but *still* often resort to low- or even no-tech, low-effort approaches. Image editing tools like Photoshop have long afforded people the ability to convincingly alter images or create new ones (Kapoor & Narayanan, 2024), and "cheap fakes" (Paris & Donovan, 2019), images taken out of context (Brennen et al., 2020), or plain false statements continue to persist and thrive. The main reason is that these basic forms of false or misleading information are 'good enough' for their intended purpose, given that they fulfill a demand for false information that is not primarily concerned with the quality of the content but rather with the narrative it supports and the political purpose it serves. In other words: High-quality false information is not even needed. Political actors can simply twist or frame true facts in a way that supports their narratives—a technique adopted with frequency, including by (political) elites. For instance, emotional (true) news stories about individual immigrants committing crimes are often instrumentalized by far right politicians to support unfounded claims about migrant crimes and ultimately to justify anti-immigration narratives (Parasie et al., 2025). Moreover, there are trade-offs between content quality and other dimensions: Higher-resolution video can sometimes look *too* polished to feel authentic, while more realistic and plausible messages may be less visually arresting—truth is often boring and—getting closer to it may inevitably make content less engaging.⁴ Again, the problem is less the improvement in content quality than the demand for misinformation justifying problematic narratives. The source of misinformation also matters more than the quality of the content (Harris, 2021): A poor quality video taken from an old smartphone shared by the BBC will be much more impactful than a high quality video shared by the median social media user. This is because people trust the BBC to authenticate the video. While GenAI can be used to increase the quality of the content, it can hardly be used to increase the perceived trustworthiness of the source.⁵

In the future, GenAI tools will continue to improve and will certainly allow for more sophisticated attempts at manipulating public opinion. While we should keep an eye out and closely monitor and regulate harmful AI uses in elections, we do not believe that improvement in the quality of AI-generated content will necessarily lead to more effective voter persuasion. Humans have been able to write fake text and tell lies since the dawn of time, but they have found ways to make communication broadly beneficial by holding each other accountable, spreading information about others' reputations, or punishing liars

Admittedly, with GenAI dramatically reducing the time and expertise needed to create high-quality outputs, this could shift the old effort-versus-reward calculus that once encouraged "cheap fakes." When a convincing deepfake now takes minutes rather than days, many propagandists might see little reason to settle for low-effort manipulations, so we should expect some crude fabrications to be replaced by AI-generated content that is both polished *and* plentiful. Whether this easier access to realism ultimately makes such content more persuasive—or, paradoxically, less believable precisely because it looks "studio perfect"—remains an open empirical question.

At the time of writing, people seem to distrust information from GenAI systems on important topics like elections. In February 2024, only 12 percent of Americans had at least some trust in information provided by ChatGPT about the 2024 U.S. presidential election, while more than a third had not heard of ChatGPT (McClain, 2024).

and rewarding good informants (Sperber et al., 2010). We expect these safeguards to hold even under conditions where content of lifelike fidelity can be created—and is available—at scale.

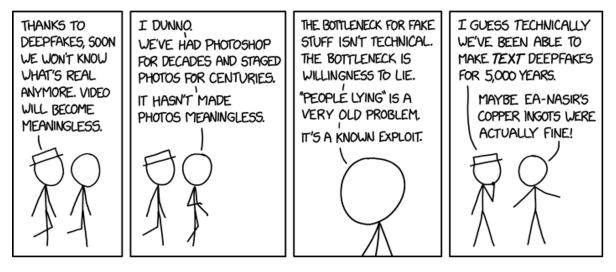


Figure 2. XKCD comic on the threat posed by deepfakes by Randall Munroe, licensed under CC BY-NC 2.5. Source: https://xkcd.com/2650/.

3. AI will improve the personalization of information and misinformation at scale

A third common argument is that AI could enhance voter persuasion by supercharging the creation of highly personalized (mis)information, including personalized advertisements. This argument is an extension of the concept of microtargeting to GenAI systems. Its popularity likely originates in the stories about the alleged outsized effectiveness of microtargeting in recent political contests, such as the supposed effects of the voter targeting efforts of political consultancy Cambridge Analytica during the 2016 European Union referendum and the 2016 U.S. presidential election (Jungherr et al., 2020; Simon, 2019). Microtargeting describes a form of online targeted content delivery (for example, as advertising or via users' in-app feeds). Users' personal data is analyzed to identify the demographic or

"interests of a specific audience or individual," based on which they are then targeted with personalized messages designed to persuade them (Information Commissioner's Office, n.d.). While addressing audiences with similar interests (e.g., for certain policy issues) and traits (e.g., age, gender) has been possible for many years and has been actively exploited by political campaigns around the world, it is far too costly to create individual messages for every single individual based on these aspects. GenAI, the argument goes, has removed this constraint, thus allowing for the effective creation of individually personalized and targeted—and thus more persuasive—content at scale.

To understand this, we must briefly examine how GenAI systems enable the creation of personalized content. GenAI systems are pretrained on a large general corpus of data and then refined in a post-training stage by approaches such as reinforcement learning with human feedback (often abbrievated as RLHF). However, current systems are unable to represent the full range of user preferences and values (Kirk et al., 2025). It is also not clear how much information current systems encode about users themselves and how much this shapes (i.e., personalizes) the responses provided to users (even though this is happening to a degree), especially around political content. To our knowledge, OpenAI's ChatGPT and Google's Gemini system are the first to incorporate a "memory" of a user's preferences and potentially use these preferences to shape subsequent responses. Beyond explicit 'memory' modules, models can also infer demographic and ideological traits from user prompts (given enough of the same) by leveraging correlations learned during pretraining. For example, experiments show that GPT-4 and similar LLMs are able to guess users' location, occupation, and other personal information from publicly available texts, e.g., Reddit posts (Staab et al., 2024). This 'latent inference' route therefore complements, and may even supersede, memory-based customization, as systems can tailor replies for people 'like you' without needing explicit profile data.

In theory, any such system that has information about an individual's traits and preferences

(or is capable of reliably inferring it) can be used to create content more aligned with the user's worldview and preferences, content that should therefore be more persuasive. Personalized information is more convincing and relevant than nontargeted information (e.g., targeted ads about local events featuring music you like, rather than generic ads) and in many ways GenAI can help personalize information, potentially making it easier to persuade or mislead people.

However, there are several complications to this claim. First, technical feasibility is not the same as actual effectiveness. The effectiveness of politically targeted advertising in general is mixed, with at best small and context-dependent effects (Jungherr et al., 2020; Simon, 2019; Zarouali et al., 2022). Previous studies on microtargeting and personalized political ads reveal that data-driven persuasion strategies often face diminishing returns without broader messaging alignment and credible, on-the-ground campaigning (Kreiss & McGregor, 2018). Experimental evidence from the U.S. further shows diminishing persuasive returns once targeting exceeds a few key attributes (Tappin et al., 2023). Skeptical voices would rightly argue that these findings do not take into account more powerful GenAI systems that can create personalized content at scale in response to customized prompts, all at little cost. However, evidence that the personalized output from AI systems is more persuasive than a generic, nontargeted persuasive message is thin. A recent review of the persuasive effects of LLMs concluded that the "current effects of persuasion are small, however, and it is unclear whether advances in model capabilities and deployment strategies will lead to large increases in effects or an imminent plateau" (Jones & Bergen, 2024, p. 25). Hackenburg and Margetts (2024a, 2024b) found that "while messages generated by GPT-4 were persuasive, in aggregate, the persuasive impact of microtargeted messages was not statistically different from that of nontargeted messages" and that "further scaling model size may not much increase the persuasiveness of static LLM-generated political messages" (Hackenburg et al., 2025). The approach used to study such questions also makes an

important difference: Studies measuring the perceived persuasiveness of (text) messages, by asking participants to rate how persuasive they find messages (e.g., Simchon et al., 2024), find much larger effect sizes than more rigorous studies that measure the actual change in participants' post-treatment attitudes (Hackenburg & Margetts, 2024a; Hackenburg et al., 2025). Similarly, the effect of political ads are much stronger when relying on self-reported measures of persuasion rather than actual persuasion, notably because people rate messages they agree with more favorably (Coppock, 2023). In general, the persuasive effects of political messages are small and likely to remain so in the future, regardless of whether they are (micro)targeted and personalized, because mass persuasion is difficult under most circumstances (Coppock, 2023; Mercier, 2020).

Second, effective message personalization plausibly requires detailed, up-to-date data about each individual. While data collection and data availability about voters is commonplace in many countries (Dommett et al., 2024; Kefford et al., 2022), including general data about traits that shape political attitudes and voting behavior (e.g., race, age, gender, partisan affiliation, voting records) this is not uniform and consistent across countries. As, for example Dommett and colleagues have found, parties in various countries routinely gather voter information via state records, canvassing, commercial purchases, polling, and now online tools, but the depth and type of data they can obtain depend heavily on each country's legal frameworks, local norms, and the specific access rules of individual jurisdictions. This data access is also shaped by the resources of political actors. Major parties generally have the resources to blend state files, purchased data sets, and digitally captured traces, whereas smaller parties often lack the funds or capacity to canvass intensively or pay for voter lists. Data about individual preferences, psychological attributes, and political views (or data associated with the same) are even more difficult to obtain. Even when userlevel data needed for precise political personalization can be obtained, hurdles remain. Data sets can be noisy and incomplete (Dommett et al., 2024; Hersh, 2015). Insufficient

temporal validity cannot be ruled out, and weak links between the data and the attempted observed construct (e.g., constructing political views from purchasing behavior) are possible. In addition, unforeseen external shocks not reflected in the data and other unobserved features not captured in the same—which would be meaningful for accurate personalization—compound these issues. These challenges, which already hobble general attempts at predicting human behavior with traditional forms of predictive AI (Narayanan & Kapoor, 2024, p. 29) will likely also complicate efforts to personalize content with GenAI systems in a way that leads to significant attitudinal or behavioral change, limiting the fidelity of any downstream personalized targeting. Again, skeptical voices will argue that the data gathered directly from voters' interactions with various AI systems (e.g., through the aforementioned memory functions) or inferred from their interactions with the same will be superior in quality. We think it is unlikely that this data would be exempt from these constraints (e.g., temporal validity, external shocks, incompleteness), or that it would become widely accessible outside of AI firms in a way that allows third parties (such as political parties and candidates or actors, or malicious actors) to easily create individually personalized messages that are then also targeted at the right person in various ways (not just within AI systems like chatbots but also in the form of ads or messages on other platforms). It is also unlikely that AI firms would allow such targeted personalization attempts within their own systems (Goldman & Kahn, 2025). Such data access and use would also be prohibited in many countries (see Dommett et al., 2024).

32

Theoretical pathways for the delivery of messages personalized with AI

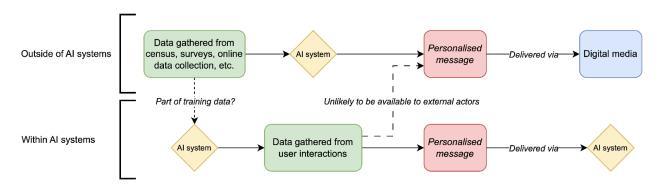


Figure 3. A schematic representation of how personalized messages could be delivered with AI.

Illustration courtesy of the authors.

Third, to be successful, a personalized message must reach the individual in question. This could happen in two ways (see Figure 3): In the first scenario, the message is delivered as part of the output of a GenAI system itself; for example, as part of a conversation with a user (we discuss this in more detail in Section 6 below). However, so far, no AI or platform company has stated an intention to provide political parties (or other actors with political intentions) with access to their systems to allow them to create and deliver targeted personalized responses at scale to individuals (as was the case, at least to a degree, with Facebook during various political campaigns in the past; Kreiss & McGregor, 2018). Any such move would likely not go unchallenged and be subject to legal restrictions in various countries. In addition, it is questionable that most users would react positively to unsolicited political messages targeted to their preferences from AI systems. In the second delivery scenario, a message is personalized using AI and information gathered about users and then delivered via existing digital platforms that they use daily. However, while GenAI systems (at least theoretically) reduce the cost of creating personalized information, it does not reduce the cost of reaching people individually via such means on these digital platforms. After all, targeting people with tailored messages online does not come for free, and certain audiences

can be more expensive to reach and prices differ between political actors (Lambrecht & Tucker, 2019; Votta, Dobber et al., 2024). In addition, as mentioned earlier, attention is a scarce resource. Any piece of politically relevant information—including personalized messages or ads, regardless of their accuracy—must compete with other types of information for peoples' attention. There is also evidence that the more people encounter ads that try to persuade them, the more skeptical they become of them and the worse they become at actually remembering the messages they have seen (Bell et al., 2022). According to political campaigners, people often do not pay attention to personalized political advertisements (Kahloon & Ramani, 2023). People also not only recognize such personalized messages but also actively dislike those that are excessively tailored (Gahn, 2024; Hersh & Schaffner, 2013) or tailored using certain traits that are considered too personal (Bon et al., 2024, Vliegenthart et al., 2024). Their use can also lead to backlash in favorability if they come from parties that voters do not already agree with (Binder et al., 2022; Chu et al., 2023; Vliegenthart et al., 2024).

Fourth, the growing abundance of voter-level data and increasingly sophisticated AI tools does not automatically ensure that political actors will deploy them. Past microtargeting efforts offer a telling example: audits of 2020 U.S. Facebook ads reveal that official campaigns used the most granular targeting mainly for highly negative messages, leaving much of the platform's segmentation potential untapped (Votta et al., 2023). While political campaigns worldwide use targeted advertising, spending is mostly "allocated towards a single targeting criterion" (such as gender) (Votta, Kruschinski, et al., 2024). While wealth-ier countries and electoral systems with proportional representation see greater amounts of money focused on microtargeting combining multiple criteria (Votta, Kruschinski, et al., 2024), European case studies document legal, budgetary, and cultural constraints that have kept microtargeting in bounds (Dobber at al., 2019; Kruschinski & Haller, 2017). Decisions in political campaigns (and in many other organizations) are also not solely, or even

primarily, driven by data or new technological systems (see e.g., Christin, 2021 for the use of metrics in news organizations); instead, human agency and organizational dynamics play a crucial role in determining when and why the same are, or are not, integrated into strategic decision-making (Dommett, personal communication, April 2025). Organizational sociology theory helps explain this gap: Rather than acting as fully rational optimizers, political actors 'satisfice' under bounded rationality and established routines. In other words: Political actors do not choose the absolute best option; instead, limited time, information, and ingrained habits lead them to pick the first solution that seems 'good enough.' For example, research on recent technology-intensive campaigns in the U.S. shows that party strategists often override model recommendations in favor of gut feeling, coalition politics, or candidate preferences (Kreiss, 2016). Reinforcing this observation, a post-2024-election report from the Democratic-aligned group Higher Ground Labs (HGL Team, 2024) observed that AI never dominated campaigning as some practitioners had predicted; most teams limited the technology to low-stakes tasks such as drafting emails, social posts, and managing event logistics, despite personalization with AI being technically feasible at the time (at least in theory). Only a handful ventured into more sophisticated uses like predictive modeling, large-scale data analysis, or real-time voter engagement, and even then, adoption was typically the result of individual experimentation rather than a structured organizational rollout. The report stresses that many staffers simply lacked the know-how to push GenAI further and found little institutional guidance to help them do so, illustrating how entrenched routines, uneven skills, and weak organizational support continue to constrain the political uptake of advanced technologies (HGL Team, 2024). A more recent survey found that while AI use among political consultants in the U.S. is growing, it is mostly used for mundane admin tasks (Greenwood, 2025). All this is unlikely to be different for techniques enabling personalized messaging with the help of generative AI systems. In short, the beliefs, capacities, and priorities of political actors remain bottlenecks between what

data and GenAI theoretically make possible in terms of personalization and what political campaigns actually do.

Fifth, and finally, the overall argument assumes that GenAI will be used to mislead more than to inform. However, AI chatbots will also be used by governments, institutions, and news organizations to inform citizens and provide them with personalized and reliable information (we discuss this further in the earlier section on mass persuasion). Personalized targeting with information is also not inherently anti-democratic; pluralist and deliberative theories of democracies hold that citizens participate most effectively when they receive information and representation that is salient to their lived interests and identities (Dahl, 1989; Mansbridge, 1999). In this sense, AI-enabled tailoring can also fit into the long-standing democratic practice whereby parties canvass different interest groups with messages that match their specific concerns and build coalitions around such issues. Such tools could in theory also expand informational equality by delivering high-quality, language-appropriate content to communities that mainstream media often underserve, including linguistic minorities, first-time voters, and rural electorates (see e.g., Vaccari & Valeriani, 2021). The democratic question is therefore less about whether personalization occurs and more about whether citizens retain exposure to diverse viewpoints and qualitative information.

4. New modes of information consumption

A fourth argument that GenAI spells trouble for elections is that the integration of GenAI systems into existing digital infrastructures such as social media and online search and the increasing use of GenAI for information seeking leads to changing consumption patterns around election information. This, in turn, could present several risks. First, that voters are more likely to be misinformed because these systems provide them with incorrect information (regardless of why they do so). Second, that they provide voters with lower-quality or biased information in the context of elections. Third, that their use crowds out or otherwise

undermines authoritative sources of information, including about elections.

As explained earlier, elections are mechanisms by which political conflict in society is channeled into real power over society. They are also a mechanism of control, allowing voters to vote out of office candidates that do not perform or do not match their preferences. Central to this is that voters are able to evaluate politicians. As Jungherr et al. write: "To keep incumbents accountable, voters need to know what politicians do; to select from among all candidates the one that will best represent them, voters need information on what politicians want" (2020, p. 216). The availability of information, and information from independent news media, are critical in this context—and a range of studies has shown a link between a better-informed public and better performance of politicians (Besley & Burgess, 2002; Brunetti & Weder, 2003; Freille et al., 2007; Snyder & Strömberg, 2010). Increasingly, GenAI is used to provide people with news, or news-like information, including about elections. At the time of writing, GenAI has been integrated into online search engines (Google, Microsoft Bing) and social media platforms (Meta's Facebook, Instagram, WhatsApp, and Messenger, as well as X), in addition to being widely available as standalone chatbots. As survey research in 2024 from Argentina, Denmark, France, Japan, the U.K., and the U.S. shows, GenAI sees increasing use, with 24 percent of respondents reporting that they used GenAI for getting information, although just five percent said they had used GenAI to get the latest news (Fletcher & Nielsen, 2024). With the growing adoption and integration of GenAI, these numbers will very likely increase over time. The latest results from the 2025 Digital News Report⁶ showed that this number has risen to seven percent at the time of writing (Newman et al., 2025).

However, GenAI, particularly chatbots, can produce plausible-looking information about elections that is partially or entirely false or misleading. For example, news outlet Proof

⁶ An annual nationally representative YouGov survey of over 92,000 news consumers in 46 markets.

News and the Institute for Advanced Study at Princeton University found in February 2024 that answers to questions about the 2024 U.S. election from five different AI models "were often inaccurate, misleading, and even downright harmful" (Angwin et al., 2024), with similar findings reported for three chatbots for the 2024 EU parliamentary elections (Marinov, 2024; Simon et al., 2024), although another study (Simon, Fletcher, & Nielsen, 2024) found somewhat better results for the 2024 U.K. general election. Meanwhile, internal research by the BBC found issues with inaccuracies and distorted content in the representation of news content, and BBC content in particular (Rahman-Jones, 2025), something also found by Jaźwińska and Chandrasekar (2025) who found that leading GenAI systems frequently misrepresented news content.⁷

Given the likelihood that more people will use such systems as sources of news, and given these systems' weaknesses, a current fear seems to be that users are more likely to be misinformed (potentially misinforming others in turn), because these systems provide them with incorrect information (regardless of why they do so), for example around elections. Particularly problematic here is that GenAI systems often only provide a single answer, which differs from search results, where users are presented with a range of options to pick from. They also can provide factually incorrect information with the same certainty as factually correct information. However, there are several reasons why the impact of this—especially around elections—might not be as severe as feared. First, receiving plausible-sounding (but potentially incorrect answers) from a GenAI system does not significantly differ from communication flows in everyday life. Statements of friends, family, or colleagues routinely include inaccuracies, yet people generally navigate these situations with ease without

⁷ It is worth noting that it is difficult to make a general statement about these systems' performance given their continuous development, the small-scale nature of these studies, and the lack of an agreed methodology for assessing the factual accuracy of such systems. However, the fact that different studies with different approaches have found broadly similar results points to a wider problem in how these models handle factual accuracy.

dramatically altering their (political) views (i.e., people are epistemically vigilant; Mercier, 2020; Sperber et al., 2010). This suggests that while a single incorrect statement generated by a chatbot might have an impact in certain contexts, it is not fundamentally different from other sources of error-laden or biased information encountered through normal social and media interactions. In addition, as we have already argued, factors such as partisan identity mediate the impact of new information voters receive on attitudes and behaviors. Second, generative systems can be—and already have been—deliberately designed to mitigate such risks. Google's Search Generated Answers, for example, do not provide answers for certain topics at all and otherwise provide users with links to sources (although it is still unclear to what extent users will interrogate these, with a growing number of accounts indicating that users often do not click through to underlying sources). Emerging models could incorporate further features that enable users to consult original, authoritative sources or weigh competing accounts, thereby reducing the likelihood of uncritical acceptance of misleading statements. They could also be designed to provide blanket template responses to particularly important questions ('When is election day?'). Third, incidental exposure to diverse information remains common across digital and offline networks (Ross Arguedas et al., 2022; Vaccarie & Valeriani, 2021), minimizing the chance that a single chatbot response will decisively shape political attitudes.

A second concern around the increasing use of GenAI for information consumption relates to the idea that it will provide users—and thus voters—with lower-quality or biased information in the context of elections. This concern does not revolve primarily around the factual accuracy of responses but issues such as the plurality of views presented (e.g.,

One wrinkle, however, is that the net negative impact could be larger if the chatbot replaces time that users would otherwise have spent with more trustworthy (news) sources; in that sense, the harm would arise less from the error itself than from the opportunity cost of crowding out higher-quality information. Then again, given the vastness of knowledge encoded into their systems and their ability to browse the web, on average LLMs are likely more accurate than some other sources of information people rely on.

voters receiving only information on the views of one party, instead of several points of view), as well as the level of detail and nuance (e.g., sensational content that oversimplifies a complex issue relevant to a given election), and the reflection of minority views in output (Jungherr, 2023). Third, and relatedly, the increasing use of AI for information and news could end up crowding out other authoritative sources of information (because they are not represented in the output) or otherwise undermine them (for example, because the creation of products competing with those of news organizations weakens their economic position), with negative downstream effects for the information available about and during elections. While none of these points have so far been empirically validated, these second-order effects could be problematic—especially for democratic life at large, given that access to quality information and news helps people be more informed and can play a role in increasing resilience to misinformation (Altay 2024; Humprecht et al., 2020)—even where they do not have major effects during elections, given how people's existing attitudes and cognitive biases shape their electoral behavior, as explained earlier.

5. Destabilization of reality

A fifth argument around AI and elections states that the ability to create realistic-looking content with GenAI will sow confusion and uncertainty about what is real, which could in turn increase skepticism toward accurate information, diminish trust in accurate information from reliable sources such as quality news outlets and experts, and be weaponized to deny inconvenient truths ('This footage is not real, it's AI!'), also known as the liar's dividend.

GenAI can indeed be used to create highly realistic images, videos, and audio. The possibility that any realistic-looking content may be AI-generated could make people more skeptical, especially when it comes from sources they do not know or trust. This heightened skepticism has not been properly empirically documented yet, although a first, small-scale experimental study in Germany showed that exposure to deepfakes significantly decreased

the general credibility attributed to all types of media, even after participants were told the content was fake. The study also found that people became less confident in their ability to discern between real and fake media, regardless of whether they were exposed to authentic or manipulated content (Weikman et al., 2024). Anecdotally, we have seen a handful of instances where people have become more skeptical—for example around images of nature on social media, or viral entertainment videos, as well as in complaints about the increase in so-called AI slop (low-quality, AI-generated content).

Yet, claims that this development will foster widespread skepticism toward all kinds of accurate information have to be taken with a pinch of salt, given the complexity of how people evaluate information. Individuals filter political information through existing mental models shaped by their political socialization and position (see the earlier background section for reference). They also rely on more than just the overt realism of a message to assess it. Instead, audiences also consider preexisting familiarity with the source and the source's credibility. As, for example, Harris (2021) argues for the case of videos, these derive their evidentiary power not solely from their quality or content "but also from its source. An audience may find even the most realistic video evidence unconvincing when it is delivered by a dubious source. At the same time, an audience may find even weak video evidence compelling so long as it is delivered by a trusted source." In addition, subtle inconsistencies in modality or content, audiences' own preexisting knowledge of a topic or claim, and the broader context in which the information is presented all play a role (Hameleers et al., 2023; Mercier, 2020). Individuals often exhibit a healthy degree of scrutiny and critical thinking when confronted with novel content, even from a young age (Harris, 2012; Sperber

It should be noted that participants in the experiment saw only one short deepfake under tightly controlled conditions and the study included no longitudinal follow-up. The post-reveal dip in credibility and self-efficacy may reflect a momentary novelty or learning effect, with no evidence yet on whether skepticism remains or deepens over time.

et al., 2010). Historical precedents such as staged photographs, heavily edited political broadcasts, or the introduction and widespread availability of Photoshop and similar editing software have likewise raised concerns about media authenticity, yet society has repeatedly developed new norms and tools to discern and counter manipulation (Habgood-Coote, 2023) and largely seems to have retained trust in depictions of reality. Moreover, ongoing advancements in digital literacy programs and AI-detection technologies could bolster the public's and experts' capacity to identify and reject manipulated media. To date, there is little evidence to suggest that the rise of GenAI has escalated into widespread doubt about the legitimacy of all information.

The suspicion and uncertainty that any content may be AI-generated, and the difficulty in some cases of detecting it, will inevitably shape norms and perceptions of information online. It is possible that general levels of skepticism and distrust in new information encountered online will rise. However, while the added noise to the information environment will be detrimental, established providers of reliable information might not see much of an effect at all. Trust in institutions and news is driven by a plethora of factors (partisanship, habituation, rituals of trustworthiness such as transparency, see Fawzi et al., 2021 for an overview), many of which are not directly affected by AI or the existence of AI content at all. On the contrary, rather than becoming universally skeptical of all content, people may become more selective about whom they trust, and gravitate even more strongly toward sources they consider authoritative or authentic. Many information channels favored by youth, such as TikTok, Snapchat, or Twitch, already promote such informal, on-the-fly, personal, even intimate content, where audiences create—sometimes strong—parasocial relationships with content creators. Doubtlessly, these sources may still disseminate misleading or false information, but this is not fundamentally an AI issue. Rather, it is a characteristic of high-choice media environments, which allow for the existence and proliferation of a plurality of sources of different quality, standards, and formats.

Meanwhile, the risk of a 'liar's dividend'—where politicians or other public figures dismiss authentic material as AI fabrications, thereby creating plausible deniability—is real. There is limited evidence that such a strategy can have some effect, for example for politicians (Schiff et al., 2023). However, it is currently unclear how effective this strategy actually is and whether AI provides a substantial advantage compared to existing techniques and technologies (including plain old lying). There are several reasons to assume that AI provides only a marginal advantage, if at all. First, politicians have always been motivated to discredit unfavorable but accurate information about themselves or their campaigns, often turning to whatever culturally available excuses happen to suit their needs, regardless of plausibility. The realism of AI certainly adds a weapon to their arsenal, but ultimately, what determines whether such attempts will be successful is not primarily the existence of more advanced technology but the receptivity of the public and people's preexisting trust in the individuals trying to discredit the evidence (Ecker et al., 2022). If convincing enough, partisans will be more likely to embrace a message of denial when it supports their party or candidates regardless of whether there is evidence to back it up. However, when the AI fabrication runs counter to people's preexisting beliefs and values, as well as readily verifiable information, they will likely remain skeptical of blanket denials. In addition, forms of digital forensics and journalistic fact-checking will likely mitigate this risk to some degree, and traditional techniques of verifying evidence remain relevant and effective, notwithstanding the existence of more powerful AI systems.¹⁰ While high-profile allegations of fabricated evidence have occasionally garnered public attention, systematic research has

To give an example, for an AI-generated image of an event, investigators can check if matching photographs or clips shot from different viewpoints, at different times, and from different sources exist and cross-corroborate it with eyewitness accounts. A range of OSINT and digital forensics approaches can be used, too, in addition to content provenance mechanisms that exist in some cases. A French investigation from May 2025 demonstrates this well: https://observers.france24.com/en/france/20250521-artificial-intelligence-detection-tools-real-image-melenchon-reliability. While none of these approaches might suffice on their own, they demonstrate that existing defenses continue to matter.

yet to demonstrate any widespread success of the 'it's just AI' defense in reshaping major political narratives.

6. Human-Al relationships

The final argument concerns potential human-AI relationships, where humans "form what they perceive as relationships with personalised and agentic AI systems capable of emotional and social behaviours" (Kirk et al., 2025). Long before the rise of GenAI, human-computer interaction researchers theorized and proved in many experiments that humans can perceive and treat computers as social actors. 11 Such relationships are theorized as deeper, more trusting, and more enduring than one-off or occasional uses of AI systems as tools. Humans view, among other things, emotional investment, mutual vulnerability, shared moral agency, humor, continuity, and existential reciprocity as foundational to meaningful relationships (Bickmore & Picard, 2005; Kirk et al., 2025; Zimmerman et al., 2024). AI agents have attained at least some of those qualities already—or are at least seen as having attained them because they engage in convincing "role play" (Shanahan et al., 2023),12 with recent research demonstrating that they "give the impression of relationship-building to human users, and that this is more likely when users interact with AI systems for high empathy, socially-oriented needs such as friendship and life coaching" (Ibrahim et al., 2025). The corollary, so the broader argument, is that this closer bond can lead to (1) an increased persuasion of users via AI systems or agents in general; (2) an increased persuasion of users through the propagation of false or biased information by the AI, given that these systems have a tendency to hallucinate or harbor in-built biases; (3) the potential for targeted manipulation of voters if malicious actors gain control over these systems; and (4)

¹¹ See, e.g., the CASA literature (Computers as social actors, 2025).

Regardless of whether they can have e.g., "moral agency" or "vulnerability" from an ontological point of view.

the reinforcement of voters' existing views if these systems are overly sycophantic and fail to challenge users' false beliefs.

Again, we address these points in turn. Specifically in response to the first argument, this point crucially hinges on two factors: (1) that AI systems in general, and especially agents that are like friends or partners, are more persuasive than humans and (2) that we will see the widespread formation of such close bonds between humans and AI agents. Starting with the latter point, while recent research and reporting acknowledge instances of intense human-AI bonding—such as users reporting emotional attachment to chatbots or virtual companions (Hill, 2025)—it is presently unclear if such deep relationships would become mainstream.¹³ While niche communities or some individuals might demonstrate intense engagement, just like some adults develop an unhealthy attachment to fictional characters (Rain & Mar, 2021), these for now seem to remain outliers and therefore a restriction on the possibility of mass persuasion from AI agents—although this number will increase with growing adoption and greater capabilities of systems across modalities and functions. Which brings us to the first point, that friend or partner-like AI systems will be more persuasive because we see them as equal to trusted friends or partners.¹⁴ There are some reasons to be skeptical of this argument. For one, from a young age, children are able to tell the difference between fiction and reality and are skilled at interacting with complex fictional worlds and characters (Harris, 2012). No matter how good virtual AI agents become, it is questionable there will be a total suspension of disbelief toward them. Recent studies also show that

Nota bene: Surveys where an increasing number of people say they have used AI systems to seek some form of companionship, emotional support, and the like are not in and of themselves proof that they have come to rely on these systems (or will do so) and that this relationship is of equal quality as a relationship to other humans.

It should be said that if this is borne out, such AI systems should not by default be presumed harmful. An AI system with whom someone has bonded that is also demonstrably more accurate on most topics than the average conversational partner may provide greater benefit and less risk of misinformation than reliance on (chance) human interactions. In short, persuasiveness grounded in genuine utility and factual reliability can be a feature, not a bug, of AI companionship.

while people can form emotionally meaningful relationships with AI chatbots (Skjuve et al., 2022), they also remain aware of their artificiality (Brandtzaeg et al., 2022), which may limit their persuasive impact.

It is possible that factors such as "immunity to fatigue" (Earp et al., 2025)—that AI systems are always available, never tired, and will always engage even when a human would not could give them a different persuasive advantage. However, there is to date no good empirical evidence of this. Perhaps the best argument against any outsized persuasive power on political matters of AI systems to which we have grown very close can be found in the literature on the one "mini-network that figures so prominently in the lives of most people" and arguably leads to some of the closest bonds humans form: romantic long-term partnerships (Jennings & Stoker, 2001). There is limited academic evidence that people significantly change their political views to match those of their romantic partners after entering a relationship, and even less showing that such change is primarily due to arguments or information exchange—in other words, persuasion. Instead, most research seems to find that "individuals with similar partisan preferences are attracted to each other because they confirm each other's worldview and identity" (Hudde & Grunow, 2025, p. 1582) and that people select partners with similar views from the outset (Easton & Holbein, 2021; Huber & Malhotra, 2017; Hudde & Grunow, 2025; Nicholson et al., 2016). While there is some evidence of a modest increase in overall ideological similarity as relationships progress (for a discussion, see Hudde & Grunow, 2025), the evidence does not clearly show that arguments or the sharing of information are the main mechanisms driving any convergence. Instead, changes—when they occur—are likely subtle, gradual, and influenced by broader relational dynamics or alignments in material conditions and incentives (e.g., income, having kids, or becoming a homeowner) rather than direct persuasion or debate. The corollary of this is: If people do not really change their political views to match those of their partners, why would we expect AI agents to be significantly more successful in this regard? Should we not

rather assume that AI systems will try to match the views of their human counterparts? This might then lead to a reinforcement effect, which could be problematic in itself but is also already very common in social life.

On the second point, while it is correct that current AI systems have a tendency to hallucinate and provide biased output, the ways humans consume information—including from friends and family members and soon personalized AI agents—and form beliefs about the world (which can shape their decisions) is marked by various epistemic or "open vigilance mechanisms" (Mercier, 2020). These include our ability to check whether a message is compatible with existing beliefs or knowledge (plausibility checking), whether it is supported by good arguments (reasoning), or comes from a reliable and well-intentioned source. In addition, information consumption does not happen in a vacuum but is embedded in structures where people are exposed to cross-cutting information. As argued previously, even a more personal relationship with an AI agent is unlikely to blast through these defenses—or if so, would not necessarily lead to a situation that is fundamentally different from the status quo, where people are already exposed to false, truncated, or misleading information from other humans with whom they have formed close bonds.

Regarding the third point, the potential for targeted manipulation of voters through agentic AI systems, we point the reader to Section 3 on the limitations of personalization and Section 4 on the use of such systems, as the same arguments apply here. Moreover, AI systems remain, to some extent, subject to existing regulations, platform policies, and security measures. The track record of digital manipulation attempts on major platforms, for example, indicates that large-scale influence operations are frequently detected and publicized—and to our knowledge, no adversary has thus far been able to take control of, e.g., Meta's algorithmic recommendation systems. We consider it, for now, unlikely that this should change, considering that AI agents are largely developed by the very same firms that usually have stringent cybersecurity measures in place, which should limit the possibility of a hostile

takeover of the same. Skeptical readers are justified in pointing out that this does not guard against the manipulation of such systems where the companies themselves engage in such activity, either in an attempt to please political actors or out of their own political motives. This is a possibility that cannot be ruled out. However, here too, regulatory frameworks and the "bulwarks" described in the sections above should present at least some friction. All but the subtlest attempts would also likely fail to go unnoticed and likely create a public and political backlash, as well as regulatory and legal interventions, against these companies.

Fourth, the concern that AI systems, by always agreeing with users, will reinforce existing beliefs and create impenetrable echo chambers stems from a deterministic view of both technology and human interaction. Research on digital media consumption shows that while selective exposure can reinforce biases, most individuals do not exist solely in sealedoff information bubbles; exposure to cross-cutting viewpoints remains common in offline networks and through incidental online encounters as well as forms of incidental news and information exposure (Beam et al., 2018; Dubois & Blank, 2018; Hartmann et al., 2024; Ross Arguedas et al., 2022). Additionally, designers of conversational agents as well as independent researchers are increasingly aware of the potential for sycophantic or manipulative behavior (Williams et al., 2025) and are exploring ways to correct for the same—by integrating strategies such as refining training data to reduce bias (Rrv et al., 2024), adjusting reinforcement learning protocols to penalize overly deferential responses (Sharma et al., 2023), and employing prompt engineering to encourage balanced dialogue (e.g., prefacing corrections with empathetic phrases like 'I understand your perspective, but research suggests...'). However, how far these efforts will be implemented and win out against a concurrent drive—fueled by economic incentives—to make AI systems into more engaging "AI companions" to keep users tuned in remains to be seen (Tiku, 2025).

The future will likely see a growing use, and in a few cases a growing dependency, on AI

systems as personal companions. There is also early evidence that AI-mediated conversations can shape user attitudes. More research will be needed to understand the risks to users stemming from those interactions, both in general and with respect to the impact on their political views and behaviors. However, we argue that there is no good reason to believe that AI companions—no matter how sophisticated—would be exempt from the complex ecosystem of cognitive, social, and institutional checks and balances that humans rely on day in, day out. The mere presence of AI capabilities and the emergence of deeper human-AI relationships does not guarantee seamless or unstoppable manipulation of public opinion or voting behavior.

Discussion

In the previous section, we discussed why common arguments that GenAI systems pose an outsized risk to elections are overstated. In this section, we will (1) consider the challenge of defining GenAI content, (2) review how people currently use GenAI content, and (3) present evidence on how GenAI was used in the 2023-2024 election cycle.

1. What counts as AI-generated content?

The actual prevalence of AI-generated content that people encounter online is an empirical question that raises fundamental definitional challenges. Understanding the scope and nature of AI-generated content is admittedly not as straightforward as it might seem. The line between what constitutes AI-generated content and what does not is difficult to draw. Social media platforms and online spaces offer a variety of AI tools that enhance or modify content, ranging from simple text-to-voice features to AI-generated images or videos. While many uses of GenAI in content creation are not problematic, distinguishing between harmless applications and those that raise ethical concerns is a complex task.

These applications of AI are often simple, incremental improvements that do not

fundamentally alter the nature of the content. In many cases, they are tools for convenience, accessibility, or creativity—helping users save time or express themselves more effectively. Importantly, most of these uses are neither misleading nor problematic. The content remains largely truthful or unchanged in its intent, even if it is enhanced by AI. However, determining where the line is crossed—where AI content becomes deceptive, manipulative, or ethically questionable—requires a better understanding of the context in which it is being used. This raises questions about the definition of AI-generated content: Does it only include content that has been completely produced by an AI model (e.g., fully AI-generated text or images), or does it also include content where AI played a supporting role in enhancing human-generated material?

In recent elections, GenAI tools have been used for what might be termed "soft propaganda" or "softfakes" (Chowdhury, 2024)—the creation of memes, imagery, and slogans (sometimes by political candidates themselves) to convey broad emotional appeals or impressions. In the 2023 Argentine general election, for example, both major candidates used AI to generate memes that either reinforced a positive image of themselves—e.g., 'I'm strong,' 'I'm a man of the people'—and a negative image of their opponents—e.g., 'He is weak,' "He's out of touch' (Nicas & Herrera, 2023). These tactics are about shaping the political mood, creating a favorable image, or invoking a visceral reaction rather than attempting to persuade or convince them through complex arguments or detailed policy proposals (the domain of "hard propaganda"). These uses—while problematic in many respects—are not new. Political cartoons, posters, and pamphlets have always aimed to simplify complex political messages into easily digestible, often misleading emotional content. AI has made this process more efficient but hasn't fundamentally changed the nature of the content.

2. How do people actually use AI?

A key consideration when evaluating AI's potential impact on elections is that most people

have limited interest in politics and news. Studies and surveys consistently show that while a vocal minority is highly engaged online, the average user consumes little civic information and rarely shares it (Nyhan et al., 2023). Some of these individuals will likely use AI to generate political memes, deepfakes, or disinformation, but the majority of people will likely use AI in mundane and benign ways—such as to summarize text, create illustrations for a presentation, for translation, or simply to have fun. Some corners of the internet are full of AI-generated content—type "baby peacock" on Google or "made it with my own hands" on Facebook. AI is already widely used to create entertaining content, and while many AI creations are disturbing, like Shrimp Jesus (literally Jesus made of shrimps) or Italian brain rot (such as Tralalero Tralala, a shark with legs and sneakers), and seem to gain traction by virtue of their weirdness, serious AI content aimed at convincing voters is much less common. An analysis of 91,452 misleading posts on X flagged through the platform's community notes system as misleading found that AI-generated misinformation was more likely to be entertainment-related (rather than health-related, for instance) and was more positive in sentiment compared to non-AI misinformation (Drolsbach & Pröllochs, 2025). A minority of people will continue to use GenAI in harmful ways, such as creating deepfakes to harass women, minorities, or political opponents. However, these behaviors reflect the individuals more than the technology itself. AI, like any tool, will be misused, but the root problem lies with those who seek to exploit it for malicious purposes. Racists and misogynists have always found ways to harm others. Similarly, while some politicians have used AI in deceptive ways—whether to smear opponents with AI-generated attack ads or to amplify misleading narratives—the underlying behavior is not new and does not require AI. Politicians have always used technology to their advantage and for their needs (such as to mobilize voters or sway them (Jungherr et al., 2020, Chapter 8), from the printing press to television to social media. GenAI offers new methods and tools for those already inclined to stretch the truth or engage in underhanded tactics. Rather than focusing exclusively on AI,

we should address the broader problem of ethical standards in politics and hold politicians accountable for their actions, regardless of the tools they use.

3. How AI was used during elections

Here, we briefly review the available evidence on the prevalence and influence of GenAI in the 2024 elections. Overall, it seems like AI was more often used for entertainment, satire, and efficiency gains rather than widespread manipulation of voters. Traditional sources of influence, such as elites and mainstream media, played a larger role than AI. In *India*, a study analyzing 2 million WhatsApp messages found that of 1,858 viral messages, fewer than two dozen contained AI-generated content, amounting to just one percent of the total (Garimella & Chauchard, 2024). Indian media entrepreneur Ritu Kapur observed that "We didn't need AI for misinformation in the Indian elections. We have plenty coming from politicians" (Leake, 2024). While AI content was prevalent during the Indian elections, some consider that the use of AI, was, overall, not problematic and even constructive: "The campaigns made extensive use of AI, including deepfake impersonations of candidates, celebrities and dead politicians. ... But, despite fears of widespread disinformation, for the most part the campaigns, candidates and activists used AI constructively in the election. They used AI for typical political activities, including mudslinging, but primarily to better connect with voters" – such as translating content in a country with more than 22 languages (Shukla & Schneier, 2024).

Around elections in the *U.K.* and the *EU*, the prevalence of viral GenAI misinformation was low according to available estimates. Research from the Alan Turing Institute found "just 16 confirmed viral cases of AI-enabled disinformation or deepfakes during the UK general election, while only 11 viral cases were identified in the EU and French elections combined. Echoing findings from our previous research, this volume was far lower than many had feared ahead of these important campaign periods" (Stockwell, 2024). This same report concludes that "there is no evidence that AI-enabled disinformation or deepfakes

meaningfully impacted UK or European election results" and that "Generative AI played less of a role in boosting the virality of disinformation compared to traditional interference methods and human influencers." However, the report also notes that AI was used in harmful ways, notably to incite hate against political figures and sow confusion. BBC journalist Marianna Spring, who tracked the use of deepfakes on social media during the U.K. election, concluded that "... in the end this wasn't a deepfake election. ... The warnings about AI were a distraction from the lack of clear solutions to problems posed by algorithms and well-practiced misinformation tactics online" (Spring, 2024).

In *Slovakia*, two days before the 2023 parliamentary elections, a fake audio clip suggesting that the pro-European candidate was committing electoral fraud went viral (Devine et al. 2024). However, the effect of this clip is unclear, and broader societal divisions, distrust in institutions, and pro-Russian sentiment within the Slovakian society likely explain both why this message gained so much traction and why the pro-Russian candidate won the election. As Nadal and Jančárik argue: "Considering these factors, the notion that Slovakia's election was the first swung by deepfakes appears reductive, fixating on the effects of the technology while overlooking the complex social, cultural, and political dynamics that propelled a pro-Russian candidate to victory. ... A narrow focus on the deepfake misses the role of public demand for Fico's message, promoting responses that address symptoms rather than root causes" (2024, p. 3).

In *Argentina*, it seems GenAI was used primarily to enhance campaign communication rather than deceive voters: "So far, the A.I.-generated content shared by the campaigns in Argentina has either been labeled A.I. generated or is so clearly fabricated that it is unlikely it would deceive even the most credulous voters" (Nicas & Herrera, 2023). In short, it was used by both candidates to speed up the creation of campaign material.

In *South Africa*, a study found that GenAI content was relatively rare, especially compared to cheap fakes: "South African voters were exposed to very little content created by GenAI.

The majority of the misleading content spread during the election was traditional mis- and disinformation, such as false headlines, allegations of voter fraud, out-of-context images, etc. There were a handful of examples of content created using GenAI technology, but the content itself was not sophisticated, consisting primarily of poorly generated videos that were easily identifiable as fabricated, spread on both open social media platforms and closed messaging platforms" (Knight, 2024).

In Bangladesh, according to one estimate, deepfakes represented only 1.9 percent of content fact-checked by professional fact-checkers, while cheap fakes were much more prevalent, with "the use of photocards and fake quotes remain[ed] predominant, and clickbait thumbnails [were] frequently used to attract the audience" (Rahman et al., 2024). In the *U.S.*, ahead of the election, the Cybersecurity and Infrastructure Security Agency (CISA) wrote that "For the 2024 election cycle, generative AI capabilities will likely not introduce new risks, but they may amplify existing risks to election infrastructure" (CISA, 2024). In September 2024, the U.S. Intelligence Community stated that "Generative AI is helping to improve and accelerate aspects of foreign influence operations but thus far the IC has not seen it revolutionize such operations" (ODNI, 2024). After the elections, it now seems clear that the impact of GenAI was not as important as imagined. As one source summarized it: "The anticipated avalanche of AI-driven misinformation never materialized. As Election Day came and went, viral misinformation played a starring role, misleading about vote counting, mail-in ballots, and voting machines. However, this chicanery leaned largely on old, familiar techniques, including text-based social media claims and video or out-of-context images" (Tuquero, 2024). The News Literacy Project, which tracked election misinformation in the U.S., found that "tricks of context—misinformation that takes an image, quote or other piece of content out of its original context in ways that change its meaning—are by far the most used misinformation tactic this election cycle. ... Only a small fraction of examples ...—about 6 percent – involve[d] the use of AI" (News Literacy Project, 2024). Finally,

a manual analysis of Meta's Ad Library and Google's Ads Transparency Center found that "there wasn't widespread use of AI in the general election to deceive voters" (Brennen et al., 2025) and that AI was more commonly used for what Matteo Wong described as "transparent propaganda, satire, and emotional outpourings" (Wong, 2024). In the U.S., the role of GenAI has been described as underwhelming (Chow, 2024). Instead, the role of influential political or politically-aligned figures such as Elon Musk, Tucker Carlson, and Joe Rogan seems to have raised greater concerns in post-election discussions.

Worldwide, on Meta platforms, the company reported that during the elections, less than one percent of all fact-checked misinformation was AI-generated (Meta, 2024). Similarly, the fact-checking organization Logically Facts reported that "AI misinformation proved far less widespread than expected, making up less than one percent of fact-checked content on Meta platforms and just 1.35 percent of Logically Facts' total of 1,695 fact-checks" (Sichova & Das, 2024). An analysis of GenAI use in elections worldwide collected by the WIRED AI Elections Project showed that "(1) half of AI use isn't deceptive, (2) deceptive content produced using AI is nevertheless cheap to replicate without AI, and (3) focusing on the demand for misinformation rather than the supply is a much more effective way to diagnose problems and identify interventions" (Kapoor & Narayanan, 2024). In some countries, AI was used for humor or parody (Spring, 2024; Stockwell, 2024). For instance, in France, AI was widely used to make viral videos of President Emmanuel Macron singing children's songs, doing makeup tutorials, or doing silly dances—videos that the president compiled and shared on TikTok to promote the Paris AI summit. In early 2025, the consensus among journalists and experts seems to be that the impact of GenAI on the 2024 elections was smaller than anticipated (Anslow, 2024; Kapoor & Narayanan, 2024; Sichova & Das, 2024; Spring, 2024;). For instance, an expert working group noted that "We noticed a change in discourse compared to the beginning of 2024, when many experts expected GenAI to disrupt the online information space around elections and to be heavily used by candidates around the world ..."

and that "At the end of our final online call, all participants admitted they initially expected GenAI to play a bigger role in the election year 2024" (Hennemann-Heldt, 2024).

We certainly need (much) better data before drawing firm conclusions. However, if the catastrophic effects initially feared had materialized, it would have left clearer traces. For example, if GenAI content had swayed voters en masse, it would have likely been detectable even with limited or imperfect data collection methods, and be reflected in polling discrepancies or observable anomalies in voter engagement metrics. Moreover, evidence that AI was not commonly used for political persuasion in 2024 necessarily constrains its potential effects. In our view, current evidence rules out strong, measurable effects of AI on elections, but does not exclude more subtle forms of AI influence—both in terms of how AI is used and the impacts it may have—which are inherently harder to detect and quantify.

GenAI content was certainly present during the 2024 elections, probably more than ever before, but its effects were not as dire as many expected. In the meantime, a handful of powerful elites are doing their best to undermine democratic institutions and pave the way to authoritarianism, with or without AI.

Round and Round it Goes? Possible Reasons for the Skewed Discourse on AI and Elections

A question we have not addressed so far is *why the discourse around AI elections has been skewed toward alarmist views*. Here, we provide an overview of the possible reasons for the skewed discourse on GenAI and elections, and the threat widely associated with the technology in this context (see Figure 4). We diagnose this alarmist discourse as an extension of what Orben (2020) has termed the "Sisyphean cycle of technology panics," with several concurrent push and pull factors within political, technological, regulatory, media, and academic communities contributing to this narrative.

The number of news stories mentioning "AI" together with "elections" and "threat" has risen since 2022, but recently declined.

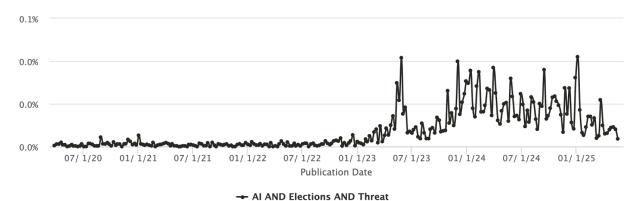


Figure 4. This chart illustrates the percentage of total coverage of stories (weekly) containing all the terms, drawn from approximately 1,500 English-speaking media sources. Source: Global English Language Sources database, provided by MediaCloud, spanning the period from April 1, 2020, to May 21, 2025.

The role of technological determinism

It makes sense to briefly consider the structural factors framing and shaping this discussion. Technological determinism—the idea that technology shapes society and is one of the most important and unstoppable drivers of change (Leonardi, 2012)——remains a popular view. This tendency to attribute excessive causal power to technology fuels both hype (Narayanan & Kapoor, 2024) and moral panic about the negative influence of new technologies, i.e., intense periods of public concern about something that is perceived to threaten societal values and social life (for a discussion, see Hoffmann et al., 2023). As Orben writes, "when a technology is first developed, marketed, and introduced, it is often either construed as good or bad for society" (Orben, 2020, p. 1146), with views frequently falling into strongly utopian or dystopian camps (Wellman, 2004). Throughout history,

We are not going to argue that technology has no effect, to say so would be wrong. We are also not going to go into an in-depth discussion of different ways of thinking about technology's role in society, which has a long and rich history. To state our own view: We are middle of the road. Technology plays a role, but not in isolation.

new technologies have been met with exaggerated hopes and fears, often framed as either transformative breakthroughs or existential threats. GenAI is no exception. After the public launch of OpenAI's ChatGPT, the salience of the topic increased dramatically (see Figures 1 and 6), with Ryauanov et al. (2025) finding that "during the six months succeeding the launch [of ChatGPT], media attention [for AI] rose tenfold—from already historically high levels." Media coverage of GenAI, however, has often been negative (Gilardi et al., 2024) and alarmist (Ryazanov et al., 2025). At the same time, recent years have seen strong concerns about the state of democracy as part of a period of, if not decline, then at least sustained democratic stagnation, and challenges to democracies around the world (Fish et al., 2018; Herre et al., 2024). Second, strong concerns about the role of technology within democracy and around elections have gained widespread traction, prominently surfacing especially after the Brexit referendum and the election of Donald Trump to the U.S. presidency in 2016 (see e.g., Jungherr et al., 2020; Karpf, 2019; Simon, 2019). The discourse around AI builds on these trends (see Figure 5). However, we identify at least six interconnected factors that have contributed to the framing of AI as a threat to elections.

The number of news articles mentioning "AI" and "elections" has grown since 2022

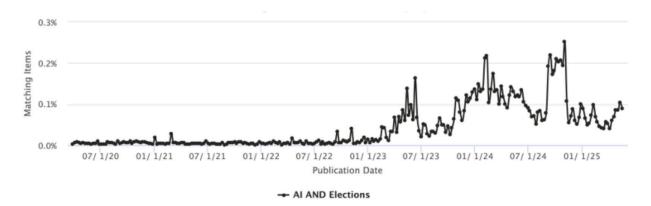


Figure 5. Percentage of total coverage of stories (weekly) containing all the terms, drawn from approximately 1,500 English-speaking media sources. Source: Global English Language Sources database, provided by MediaCloud, spanning the period from April 1, 2020, to May 21, 2025.

The number of news articles mentioning an AI-related term has grown since 2022 while the number of news articles mentioning climate change has declined.

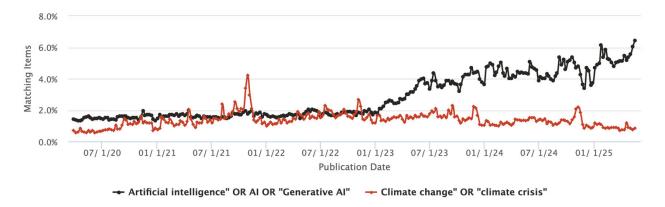


Figure 6. Percentage of total coverage of stories (weekly) containing all the terms, drawn from approximately 1,500 English-speaking media sources. Source: Global English Language Sources database, provided by MediaCloud, spanning the period from April 1, 2020, to May 21, 2025.

Self-serving claims by industry, academia, and experts

First, there are *strong*, *at times self-serving claims by industry, academia, and experts about the power of this new technology*, mixed with genuine advancements in capabilities (Narayanan & Kapoor, 2024) that have not been fully explored and therefore leave room for uncertainty. Different actors have different motivations here. For example, parts of the technology industry have strong financial incentives to support (or at least not disavow) arguments about the prowess of their technology to attract new funding and new customers—in much the same way that earlier claims about the outsized power of platforms to influence users in profound and measurable ways has likely only made them more attractive in the eyes of advertisers (Bernstein, 2021). Claims about the negative role technology might play in democracy, particularly during elections, nevertheless serve to reinforce the argument that the technology has almost magical functionality and effects.

Competing levels of expertise

Second, there are also competing and differing levels of expertise, with computer scientists

arguing with economists arguing with political scientists (and so forth) and collectively everyone else. However, given that AI these days largely emanates from the computer sciences and the technology industry (Ahmed et al., 2023; The Economist, 2023), these voices have assumed primacy in public discourse (often because they are treated by other elite actors as "high priest" equivalents, a privileged caste with better access and insight into the technology and therefore a presumably better understanding). This, however, is problematic. Computer scientists working on AI are not necessarily the best experts on the societal impact of technology, nor are they the best or only experts on (artificial) intelligence. In addition, the quarrel between different disciplines and lack of collaboration can lead to mismatched concerns, with each side prioritizing different risks and benefits based on their perspective and expertise.

Dynamics of the attention economy

Third, the *dynamics of the attention economy* matter. As Williams (2022) has argued, we live in a "marketplace of rationalisations" where "ambitious individuals and firms compete to produce intellectual ammunition for society's political and cultural factions. In return for their often-intense cognitive labor, the winners of such competition receive attention, status, and financial rewards." GenAI, of course, is not exempt from these dynamics. However, the competition for attention is not always governed by rational reasoning and deliberation about the best available evidence (and the limitations thereof). Instead, messages containing strong, assertive, or sensational claims tend to attract more attention than more

As, for example, leading AI researcher Melanie Mitchell argues: "Journalists often interview AI people but not developmental psychologists or cognitive scientists who study biological intelligence—for example, people who think about animal intelligence and how to evaluate it. ... There's no reason why AI researchers should be the only ones we hear from about the nature of intelligence" (Mitchell, as cited in Borchardt et al., 2024, 165).

moderate or cautiously phrased statements. ¹⁷ Those who make bold claims often receive more attention, which in turn can lead to further media coverage, more attention from policymakers, and increased funding, thus cementing some views over others in the public sphere. All these benefits provide further incentives to make such claims, instead of more nuanced arguments. The 'shiny new things' syndrome also comes into play: it is often more entertaining to think about the big questions (AI will end elections as we know them) than about the more mundane but often more consequential aspect of (dys-)functional democratic life (the intimidation of election workers). The problem is that early and vocal actors can shape the narrative, overshadowing or crowding out more measured views. Meanwhile, the demand from funders, policymakers, the media, and the public for answers and more information on phenomena relating to GenAI, especially with a view to its broader implications for various parts of life, can further intensify these dynamics.

Political opportunism

Fourth, *technological changes present challenges and opportunities to political actors*. There exists a bias toward action when new technology emerges. Politicians do not want to be seen as passive and doing nothing, especially in the context of technological upheavals of the past and public demand for calls to intervene. However, apart from helping them to shape their public image, moral panic around new technologies, we submit, can also be used to deflect harder conversations about social reform, justice, equality, and economic

For example, large-scale investigations of information diffusion on social media have found that content featuring novel and emotionally charged claims is more likely to be noticed and rapidly shared—suggesting that "strong claims" cut through the noise of abundant information, at least in their initial appeal. One study, for instance, showed that when users were simply "dwelling" on posts, those containing more sensational content captured longer attention spans before viewers moved on, implying that the intensity or extremity of a claim can drive initial engagement. Furthermore, modeling of collective attention dynamics supports the idea that in a saturated information environment, messages with strong claims are more likely to spark rapid bursts of public interest (see Epstein et al., 2022; Lorenz-Spreen et al., 2019).

Although the picture on the latter is complex and varies from country to country, see, for example Vogels (2022) for attitudes in the U.S. and Ejaz et al. (2024) for an international comparison.

opportunity, and steer public attention away from more intractable and uncomfortable issues. In addition, narratives of all-powerful technology endangering elections and democracy also fit well into existing arguments that call for stronger regulations of technology companies on various other grounds (e.g., data abuses, outsized market power, etc.). While we make no comment on this debate here as it is beyond the scope, AI's supposed outsized threat to the integrity of elections has been used by some political actors—but also activists, journalists, and academics—in this vein.

The window of opportunity effect

Fifth, and related to the fourth point, is what we term the *window of opportunity effect*: the sense that action has to happen immediately before a new technology calcifies. Again, there are good reasons to operate under this maxim. As historian of technology Thomas P. Hughes has argued, large technological systems (such as AI) have a tendency to become more entrenched and rigid as time goes by, making it easier to shape their direction at the beginning than during later stages (see Hughes, 2012). Applied to AI and elections, it is reasonable that the awareness of such dynamics leads to calls to safeguard against potential negative effects before they can materialize.

Genuine concern and personal experience

Sixth, and finally, we argue that *genuine concern and personal experience* of those operating in already difficult conditions or unstable political environments partially drive the one-sided debate around AI and elections. Journalists, politicians, academics, and civil society actors in countries around the world have seen digital technologies being used to cause harm in the context of elections (and beyond), for instance as surveillance structures or tools for inciting and spreading hate. For those with personal experience fighting against the recession of—or in the worst case, for the survival of—democracy in their countries, as well as those bearing the brunt of trying to uphold an epistemic order focused on truth-seeking, the open exchange of ideas and information, and the respect

of institutionalized and certified knowledge—such as public authorities and experts, scientists, and journalists—will rightly have a different perspective on how technological advancements can make their lives and work harder than it already is. This may lead to a different assessment of the impact of AI on elections compared to those not exposed to the same conditions.

Consequences of a Skewed AI Discourse

Skewed discourses do not happen in a vacuum and are not without consequences. Apart from reheating old arguments about the power of technology that have been bubbling away since the dawn of time, there are risks that flow from such a situation. In the following section, we outline what we see as the most problematic consequences.

The focus on misuses of GenAI in elections can divert us from other harms

By overemphasizing the risks of GenAI in the context of elections, we risk overlooking the broader, more insidious ways in which GenAI is misused, such as enabling targeted harassment and amplifying harmful biases. These include the *harassment of women and minorities*. The creation and distribution of AI-generated fake nudes, mostly targeted at females, is a form of gendered violence that seeks to silence women in public life (Murgia, 2024) and can be used to humiliate, discredit, and threaten women, which may have a chilling effect on their participation in politics. Similarly, minorities are targeted by AI-assisted harassment campaigns, including racially biased or xenophobic attacks that are amplified through social media. These targeted campaigns undermine efforts to build inclusive political spaces.

An overly alarmist focus on GenAI risks obscuring equally critical, long-standing threats to electoral integrity. A wealth of research in political science shows that free and fair elections depend on a complex set of structural conditions and procedural safeguards (Alihodžić et al., 2024; Norris, 2015). Any violation of these likely carries greater risks than any

that AI can currently bring about. This starts with unfit or unfair electoral systems that fail to ensure broad representation, transparency, and accountability. The poor regulation of campaign finance laws can skew the competition between political actors and give some of them an unfair advantage (Falguera et al., 2014). Fair elections also require that electoral management bodies remain independent, impartial, and transparent; where design flaws persist or where these bodies are deliberately undermined, hollowed out, or abolished, electoral disputes can escalate and erode trust in elections (Elklit & Reynolds, 2005). Mechanisms for adjudicating electoral disputes are likewise important, as slow or biased processes can undermine public confidence in the outcomes of elections (Kelley, 2012). At an operational level, logistical shortcomings and failures such as mishandled voter registrations, ballot (re-)counting, or complex and unfair special voting arrangements can be problematic if they unfairly put some voters at a disadvantage. Then there are attempts to limit ballot access—through restrictive voter ID laws, systematic purges of voter rolls, gerrymandering, and intimidation of voters—and the manipulation of electoral governance mechanisms, which undermines elections in ways that likely outweigh the existing as well as the as-yet-unrealized risks posed by GenAI (Bermeo, 2016; Ginsburg & Huq, 2018; Norris, 2015). Another pressing concern for both broader democracy and the integrity of elections lies in the systematic curtailment or weakening of press freedom and freedom of expression through various avenues. Legal and extralegal measures—from harsh libel laws to forms of media capture or outright violence or threats against journalists—can silence critical reporting and stifle dissenting voices, undermining one accountability mechanism in democracies. It can also lead to a poorer representation of the diverse elements of society and a worse voicing of grievances in society, thus impeding citizens' ability to make better-informed electoral choices.

Narratives about the outsized effect of AI on elections could lead to suboptimal policy responses

Following on from the previous section, an overemphasis on AI as a threat to elections may lead to some suboptimal or even counterproductive policy responses. Excessive or overly broad regulations could not only be ineffective—because they end up targeting the wrong thing without doing much to improve actual threats to elections and democracy—they could also backfire. In an attempt to curb, for example, AI-driven misinformation, manipulation campaigns, and the like, governments might implement sweeping measures that (inadvertently) limit free speech or restrict access to information. This could create a chilling effect, where legitimate political discourse and dissent are stifled, thereby undermining democratic principles (Center for News, Technology & Innovation, 2024a, 2024b). It also sits uneasily with the fact that governments themselves are using AI for propaganda campaigns (with, for example, the U.S. Department of Defense in the past exploring the creation of fake online personas; Biddle, 2024).

The alarmist discourse on the effect of AI on elections could reduce trust and satisfaction with democracy

The narrative that AI poses an outsized threat to elections may, in itself, contribute to a decline in public trust and confidence in democratic institutions (Jungherr & Rauchfleisch, 2024). The perception, partially co-created by media coverage, that AI has significant effects on elections could diminish trust in democratic processes and weaken the public's acceptance of election results. A recent study on concerns around the use of AI during the 2024 U.S. presidential election and public perceptions of AI-driven misinformation found that four out of five respondents expressed some level of worry about AI's role in election misinformation, with higher consumption of AI-related news linked to heightened concerns about AI's role in election information (Yan et al., 2025).

When media and public discourse emphasize the risks posed by AI, they may create a sense

of inevitability that election outcomes are manipulated or influenced by AI, *eroding trust in electoral systems and processes*. Even if AI does not play a major role in a specific election, the mere perception that it could have corrupted the process could lead voters to doubt the legitimacy of the results. This perception could be particularly harmful in tightly contested elections, where public trust in the outcome is essential to maintaining political stability. The same could also lead to a decrease in voter engagement and satisfaction with democracy. If voters believe AI is undermining the integrity of elections, they may become more disillusioned and disengaged from the political process in general. Why participate in an election, after all, if the outcome is rigged?

Relatedly, a risk in this context is the third-person effect and declining trust in one's fellow citizens: AI-generated content may shape people's perceptions and attitudes not because they themselves are influenced, but because they believe others are. This argument has been made—and empirically tested—regarding other forms of influence, such as disinformation campaigns, propaganda, and misinformation (Huang & Cruz, 2021). A U.S. study on how the general US public perceives and reacts to ChatGPT found that "individuals tend to believe they would personally benefit from the positive influence of ChatGPT, while others will benefit relatively less" and that they would be "more capable of using ChatGPT critically, ethically, and efficiently than others" (Chung et al., 2025), hinting at the possibility that the third-person effect may extend to AI, too.

Another effect may be a *heightened skepticism*: Alarmist narratives may encourage the belief that AI has infiltrated or compromised the credibility of trustworthy media and information sources. As a result, the public may increasingly question the reliability of the information they consume—although it is unclear how this would affect trust in broadly relied-upon sources of information like mainstream media. For instance, alarmist media coverage of misinformation has mostly framed the issue as a "social media problem," and while exposure to such coverage diminishes trust in news on social media, it increases trust

in print news (Thorson, 2024).

Two final points need to be made here. First, we do not want to be alarmist about how narratives on the influence of AI affect elections, either. As with the effects of AI, these are going to be small in comparison to other structural effects that have led to a steady decline in trust in institutions over the last decades (Valgarðsson et al., 2025) and in news media (Fletcher et al., 2025). Second, we do not want to dismiss voices that have raised the alarm about the effect of AI on elections. Even if we are correct that concerns about the influence of AI on elections have been exaggerated, it is possible that these concerns may have been beneficial and contributed to the current state of affairs. For example, the at times ill-informed discussions of an "infodemic" during the COVID-19 pandemic created a trading zone for different stakeholders to meaningfully engage around a shared set of concerns and problems, despite the weak theoretical and empirical evidence for the existence of such an infodemic (Altay et al., 2022b; Simon & Camargo, 2021).

In a similar vein, there is good to be found in being vigilant about the risks posed by GenAI, collectively adapting to the novel challenges the technology brings, holding the companies developing these technologies—often with little oversight and accountability—accountable, and minimizing the harm that can be caused by the deployment and use of GenAI systems. It is entirely possible that the costs of overreacting to the risks posed by AI are on average lower than the costs of underreacting in the long run. But we should keep in mind the opportunity costs of focusing on AI rather than other causes of democratic dysfunction, and the policies that will follow from focusing on AI, how these policies may be instrumentalized, and the broader effects of alarmist narratives about AI.

Conclusion

In this article, we express skepticism about the negative effects of GenAI on elections in broadly democratic countries. We do not discuss possible beneficial uses of AI to inform

voters about elections or politicians' positions or to improve democratic discussions or assemblies. One could argue that the negative effects of AI will be counterbalanced by its positive effects. We do not fully subscribe to this argument and consider that the positive effects of AI on elections and democracy more broadly are still very speculative at this stage. For instance, during the 2024 elections, there is little evidence that AI had any tangible positive effects on democracy (although see Kapoor & Narayanan, 2023 for a list of beneficial uses, or Shukla & Schneier, 2024). GenAI will be used both to promote as well as harm democracy, and in both cases, these effects will likely be smaller than expected and dwarfed by other, structural factors that affect the nature of political change and contest in democratic societies.

Some might contend that the arguments advanced here overlook significant contextual differences, since they draw heavily on examples from democracies and, to a lesser extent, non-WEIRD (Western, Educated, Industrialized, Rich, and Democratic) societies. This is correct. However, it is likewise important to avoid a paternalistic assumption that populations in the Global South or in WEIRD countries are necessarily more vulnerable to misinformation or the effect of GenAI on a cognitive level, although institutional capacities that provide a counterweight (e.g., a free press) are often less developed. Furthermore, some will rightly point out that the situation we describe could be different in autocracies and hybrid regimes. We agree, but in contexts where democracy is no longer the overarching paradigm and "pseudo-elections" are the norm, GenAI does not need to make a difference, as the outcome is already largely predetermined (in countries where no elections are held, our arguments do not apply anyway). Rather than from AI, the headline threats to political choice still come from familiar instruments: Media and internet blackouts, surveillance, intimidation, and political violence present the bigger challenge in such contexts. Moreover, we should again not assume that people in authoritarian countries are more gullible or have no way of resisting and interrogating messages they encounter for their truthfulness.

Nevertheless, caution is necessary, and dismissing AI outright could be premature, with the affordances of AI amplifying the very offline tools and strategies that make outcomes predetermined in the first place. Because the empirical record on this front is thin, we echo calls in practitioner consultations and scholarship for systematic research that pays closer attention to the effects of AI use in autocracies and hybrid regimes.

Finally, we do not intend for this article to be taken as evidence that AI poses no risks at all, that policy responses are unnecessary, or that firms developing AI should receive carte blanche. We would also like to caution against either minimizing or magnifying those risks on the basis of what is, at present, still a thin empirical record in some respect, Brennen et al. (2025) have rightly pointed out. While we think that the existing empirical and theoretical evidence shows that GenAI has not played a significant role in past elections and is unlikely to have outsized effects in the future, shrinking or nonexisting access for independent researchers to data from digital platforms and AI companies complicates researchers' capacity to study these phenomena (Mimizuka et al., 2025). Nevertheless, our position underscores the need for further empirical research across different political and cultural contexts. Below, we outline a research agenda to encourage a deeper and ongoing investigation into AI's evolving role in electoral processes.

Research Agenda

Table 2. Research agenda for the role of AI in elections

Argument	Explanation of claim	Op	Open Questions	
1. Increased quantity	Due to its technical	1.	What is the prevalence, modality, and	
of information and	capabilities and ease of use,		aim of Al-generated information in	
misinformation	GenAl can be used to create		online spaces relevant to elections?	
	information at scale with	2.	How does the volume of AI-generated	
	great ease.		content compare to human-generated	
			content?	
		3.	What are the sources and distribution	
			networks for AI-generated election	
			information?	
		4.	How does Al-generated content affect	
			the tone and civility of online political	
			discussions?	
2. Increased quality	Due to its technical	1.	How do people perceive the quality	
of information and	capabilities and ease of		of Al-generated content compared to	
misinformation	use, GenAl can be used		human-generated content?	
	to create information or	2.	How does the perceived quality of AI-	
	misinformation perceived		generated content affect its impact?	
	to be of high quality at low			
	cost.			
3. Increased	Due to its technical	1.	How effective are personalized AI-	
personalization of	capabilities and ease of use,		generated messages in influencing	
information and	GenAl can be used to create		political attitudes and behavior?	
misinformation at	high-quality (mis)information	2.	Do personalized messages have a	
scale	at scale personalized		greater impact than generic messages?	
	to a user's tastes and	3.	How does personalization interact	
	preferences.		with other factors, such as preexisting	
			beliefs?	

4. New modes	The integration of GenAl	1.	How are people using GenAl systems
of information	into existing digital	"	to access election information?
consumption	infrastructures and the	2.	How do these new modes of
Consumption	increasing use of GenAl for	۷.	consumption affect people's
	_		
	information seeking leads		understanding of election issues and
	to changing consumption		candidates?
	patterns around election	3.	How accurate are GenAl systems (on
	information.		political questions)? How accurate
			are GenAl systems in comparison to
			average and expert humans?
		4.	How does the information provided by
			AI systems perform in terms of detail,
			plurality of views, and overall quality?
		5.	Does the use of GenAI systems lead to
			substitution effects of other media?
		6.	Which providers of information get
			(dis-)empowered by the existence of
			GenAl systems?
5. Destabilization of	The realism of GenAI	1.	How does the realism of Al-generated
reality	content creates uncertainty		content affect people's trust in
	regarding what is real.		information sources?
		2.	Does exposure to realistic Al-generated
			misinformation lead to increased
			political cynicism or distrust?
		3.	How can individuals and institutions
			develop resilience and new practices
			and norms around synthetic media?

6. Human-Al	Users form deeper, more	1.	Do users form deeper relationships
relationships	persistent relationships with		with AI systems at scale?
	personalized and agentic AI	2.	Are deeper relationships with AI
	systems.		systems sufficient for short-term and
			long-term (political) persuasion?
		3.	Which users are most at risk?
		4.	Does attachment to AI systems
			(including as information sources)
			make individuals more susceptible to
			manipulation?
General		1.	How are political actors using GenAl?
		2.	Where does GenAI make a meaningful
			difference to the fulfillment of political
			actors' goals and needs?
		3.	What strategies do individuals and
			institutions use to identify and resist
			Al-generated misinformation?
		4.	How can AI be used to improve the
			quality of political discussions and
			debates?
		5.	What regulatory frameworks are
			needed to mitigate the risks of AI in
			elections while preserving its potential
			benefits?

Coda: Technological shaping is complex

Technologies and technological systems such as the internet—and now GenAI—influence how we communicate, interact, and process information about politics, but it is equally important to recognize that people have agency over how they use and adapt these technologies (Schroeder, 2018). The concept of technological determinism—the idea that technology directly drives social and cultural change in a linear, inevitable manner—has long

been abandoned. Scholars now recognize that technology's impact is complex and shaped by the ways individuals, communities, and societies interact with and tame these tools and systems in unexpected ways. Technologies do not shape society in a vacuum; instead, people interpret, repurpose, and even resist technologies based on cultural, political, and social contexts. The influence of technology is contingent, negotiated, and far more unpredictable than deterministic frameworks would suggest.

In the case of GenAI and its role in national elections and political communication, this shift in perspective is crucial. While AI systems can shape the information environment, the infrastructure of communication, and the way people receive and consume information about elections, they do not do so in a straightforward or uncontested way. People use AI for many purposes—some beneficial, some harmful—but the technology itself only dictates how it will be used to a certain extent. The impact of AI on elections is shaped just as much by human agency, creativity, and regulation as by the technical capabilities of AI itself. Many of the concerns about the potential impact of GenAI on elections and the broader information environment echo earlier panics about the influence of past technologies. For instance, when radio first became widespread, there were fears that it would centralize information and lead to mass manipulation of public opinion. Later, similar concerns were raised about the visual manipulation capabilities of software like Photoshop, which some believed would destroy the public's ability to distinguish between real and fake images despite the fact that media "history shows that evidence does not speak for itself" and that "media requires social work for it to be considered as evidence" (Paris & Donovan, 2019, p. 22)

History shows that society adapts to these technologies. Just as people learned to question the authenticity of photos and developed media literacy skills to navigate a world with manipulated images, they will likely adapt to the presence of AI-generated content and persuasive AI systems. The initial panic subsides as people develop strategies to mitigate the

risks and regulate the use of the technology. What seemed like an existential threat to information integrity in earlier decades—whether through radio propaganda or photo manipulation—eventually became manageable challenges that were addressed through education, regulation, social norms, and innovation. We argue that the same will be the case for GenAI and its role in elections.

Acknowledgments

We owe a special debt to Keegan McBride and Hugo Mercier, whose guidance in the project's earliest stage helped crystallize many of its central themes and ideas. Charlotte Jee helped us sharpen our early arguments in a piece for the MIT Technology Review, out of which this manuscript grew. We are very grateful for the support of the Knight First Amendment Institute at Columbia University, and in particular to Seth Lazar and Katy Glenn Bass, who invited us to join their "Artificial Intelligence and Democratic Freedoms" initiative and offered incisive feedback throughout. Joshua P. Darr provided invaluable comments on the first full draft, while an anonymous reviewer sharpened the second. The initial version also benefited from close readings and comments by Michael H. Tessler, Hoda Heidari, Ranjit Singh, Borhane Blili-Hamelin, Max Nickel, Manon Revel, Tyler Lu, Smitha Milli, Luke Thorburn, Henry Farrell, Atossa Kazirsadeh, and Rakshit Trivedi. A later draft was further strengthened through detailed comments and suggestions from Kate Dommett, Fabrizio Gilardi, Kobe Hackenburg, Lujain Ibrahim, Ralph Schroeder, Jan Rau, Michelle Disser, Waqas Ejaz, Tali Aharoni, Sam Gregory, Hugo Mercier, and conversations with Prathm Juneja, Hannah Kirk and Christopher Summerfield. The final version benefited from the very thorough work of two anonymous copy editors. Any remaining errors are, of course, our own.

References

- Ahmed, N., Wahed, M., & Thompson, N. C. (2023). The growing influence of industry in AI research. *Science*, 379(6635), 884–886. https://doi.org/10.1126/science.ade2420
- Aggarwal, M., Allen, J., Coppock, A., Frankowski, D., Messing, S., Zhang, K., Barnes, J, Beasley, A., Hantman, H., & Zheng, S. (2023). A 2 million-person, campaign-wide field experiment shows how digital advertising affects voter turnout. *Nature Human Behaviour*, 7(3), 332–341.
- Allcott, H., Gentzkow, M., Levy, R. E., Crespo-Tenorio, A., Dumas, N., Mason, W., Moehler, D., Barbera, P., Brown, T. W., Cisneros, J. C., Dimmery, D., Freelon, D., González-Bailón, S., Guess, A. M., Kim, Y. M., Lazer, D., Malhotra, N., Nair-Desai, S., Nyhan, B., ... & Tucker, J. A. (2025). The effects of political advertising on Facebook and Instagram before the 2020 US election (No. w33818). National Bureau of Economic Research.
- Alihodžić, S., Asplund, E., Bicu, I., & Wolf, P. (2024, September 10). *Electoral risks: Guide on internal risk factors (3rd ed.)*. International Institute for Democracy and Electoral Assistance. https://doi.org/10.31752/idea.2024.40
- Altay, S., Berriche, M., & Acerbi, A. (2023). Misinformation on misinformation: Conceptual and methodological challenges. *Social Media + Society*, *9*(1), 20563051221150412.
- Altay, S., & Acerbi, A. (2024). People believe misinformation is a threat because they assume others are gullible. *New Media & Society*, 26(11), 6440–6461.
- Altay, S., Hacquin, A. S., & Mercier, H. (2022a). Why do so few people share fake news? It hurts their reputation. *New Media & Society*, 24(6), 1303–1324.
- Altay, S., Nielsen, R. K., & Fletcher, R. (2024). News can help! The impact of news media and digital platforms on awareness of and

- belief in misinformation. *The International Journal of Press/Politics*, 29(2), 459–484. https://doi.org/10.1177/19401612221148981
- Altay, S., Nielsen, R. K., & Fletcher, R. (2022b). Quantifying the "infodemic": People turned to trustworthy news outlets during the 2020 coronavirus pandemic. *Journal of Quantitative Description: Digital Media*. https://doi.org/10.51685/jqd.2022.020
- Altay, S., de Araujo, E., & Mercier, H. (2021). "If this account is true, it is most enormously wonderful": Interestingness-if-true and the sharing of true and false news. *Digital Journalism*, 10(3), 373–394. https://doi.org/10.1080/21670811.2021.1941163
- Angwin, J., Nelson, A., & Palta, R. (2024, February 27). *Seeking reliable election information? Don't trust AI*. Proof News. https://www.proofnews.org/seeking-election-information-dont-trust-ai/
- Anslow, L. (2024). *Deepfakes haven't destroyed democracy yet*. Daily Beast. https://www.thedailybeast.com/deepfakes-havent-destroyed-democracyyet/
- Arceneaux, K., & Johnson, M. (2013). Changing minds or changing channels?: Partisan news in an age of choice. *University of Chicago Press*.
- Arno, A., Thomas, S. (2016). The efficacy of nudge theory strategies in influencing adult dietary behaviour: A systematic review and meta-analysis. *BMC Public Health 16*, 676. https://doi.org/10.1186/s12889-016-3272-x
- Aspen Digital. (2024, April 2). *AI election threat is significant*. AI Elections Initiative. https://aielections.aspendigital.org/blog/ai-election-threat/
- Bader, P., Boisclair, D., & Ferrence, R. (2011). Effects of tobacco taxation and pricing on smoking behavior in high risk populations: A knowledge synthesis. *International*

- Journal of Environmental Research and Public Health, 8(11), 4118–4139.
- Bailey, P. E., Leon, T., Ebner, N. C., Moustafa, A. A., & Weidemann, G. (2023). A meta-analysis of the weight of advice in decision-making. *Current Psychology*, *42*(28), 24516–24541.
- Beam, M. A., Hutchens, M. J., & Hmielowski, J. D. (2018). Facebook news and (de)polarization: Reinforcing spirals in the 2016 US election. *Information, Communication & Society,* 21(7), 940–958. https://doi.org/10.1080/1369118X.2018.1444783
- Becker, K. B., Simon, F. M., & Crum, C. (2025). policies in parallel? a comparative study of journalistic AI policies in 52 global news organisations. *Digital Journalism*, *o*(0), 1–21. https://doi.org/10.1080/21670811.2024.2431 519
- Bengio, Y., Mindermann, S., Privitera, D., Besiroglu, T., Bommasani, R., Casper, S., Choi, Y., Fox, P., Garfinkel, B., Goldfarb, D., Heidari, H., Ho, A., Kapoor, S., Khalatbari, L., Longpre, S., Manning, S., Mavroudis, V., Mazeika, M., Michael, J., ... Zeng, Y. (2025). *International AI Safety Report*. arXiv. https://arxiv.org/abs/2501.17805
- Bell, R., Mieth, L., & Buchner, A. (2022). Coping with high advertising exposure: A source-monitoring perspective. *Cognitive Research: Principles and Implications*, 7(82), 1–19. https://doi.org/10.1186/s41235-022-00433-2
- Bell, E. (2023, March 3). Fake news, ChatGPT, truth, journalism, disinformation. The Guardian. https://www.theguardian.com/commentisfree/2023/mar/03/fake-news-chatgpt-truthjournalism-disinformation
- Benson, T. (2023, August 1). This disinformation is just for you. Wired. https://www.wired.com/story/ generative-ai-custom-disinformation/

- Bermeo, N. (2016). On democratic backsliding. *Journal of Democracy*, 27(1), 5–19.
- Bernstein, J. (2021, September). *Bad news: Selling the story of disinformation*. Harper's Magazine. https://harpers.org/archive/2021/09/bad-news-selling-the-story-of-disinformation/
- Besley, T., & Burgess, R. (2002). The Political Economy of Government Responsiveness: Theory and Evidence from India. *The Quarterly Journal of Economics*, 117(4), 1415–1451. http://www.jstor.org/stable/4132482
- Bickmore, T. W., & Picard, R. W. (2005).

 Establishing and maintaining longterm human-computer relationships.

 ACM Transactions on Computer–Human Interaction, 12(2), 293–327. https://doi.org/10.1145/1067860.1067867
- Biddle, S. (2024, October 17). *Penta-gon seeks to use deepfakes for online influence campaigns*. The Intercept. https://theintercept.com/2024/10/17/pentagon-ai-deepfake-internet-users/
- Binder, A., Stubenvoll, M., Hirsch, M., & Matthes, J. (2022). Why am I getting this ad? How the degree of targeting disclosures and political fit affect persuasion knowledge, party evaluation, and online privacy behaviors. *Journal of Advertising*, *51*(2), 206–222. https://doi.org/10.1080/00913367.2021.2015
- Bon, E., Dommett, K., Gibson, R., Kruikemeier, S., & Lecheler, S. (2024). Are certain types of microtargeting more acceptable? comparing us, German, and Dutch citizens' attitudes. *Media and Communication*, 12, Article 8520. https://doi.org/10.17645/mac.8520
- Borchardt, A., Simon, F. M., Zachrison, O., Bremme, K., Kurczabinska, J., & Mulhall, E. (2024). *Trusted journalism in the age of generative AI*. EBU. https://www.ebu.ch/files/live/sites/ebu/files/Publications/Reports/open/News_report_2024.pdf

77

- boyd, d., & Golebiewski, M. (2018, May). *Data voids: Where missing data can easily be exploited*. Data & Society. https://datasociety.net/wp-content/uploads/2018/05/Data_Society_Data_Voids_Final_3.pdf
- Brady, H. E., Verba, S., & Schlozman, K. L. (1995). Beyond SES: A resource model of political participation. *American Political Science Review*, 89(2), 271–294.
- Brandtzaeg, P. B., Skjuve, M., & Følstad, A. (2022). My AI friend: How users of a social chatbot understand their human–AI friendship. *Human Communication Research*, 48(3), 404–429. https://doi.org/10.1093/hcr/hqacoo8
- Brennan Center for Justice. (2024, December 19). Online ad spending in the 2024 election topped \$1.35 billion. https://www.brennancenter.org/our-work/analysis-opinion/online-ad-spending-2024-election-topped-135-billion
- Brennen, S. B., de la Puerta, C., & Sanderson, Z. (2025, March 4). When it comes to understanding AI's impact on elections, we're still working in the dark. Center for Social Media and Politics. https://csmapnyu.org/impact/policy/when-it-comes-to-understanding-ais-impact-on-elections-were-still-working-in-the-dark
- Brennen, J. S., Simon, F. M., & Nielsen, R. K. (2020). Beyond (mis)representation: Visuals in COVID-19 misinformation. *The International Journal of Press/Politics*, 26(1), 277–299. https://doi.org/10.1177/1940161220964780
- Brunetti, A., & Weder, B. (2003). A free press is bad news for corruption. *Journal of Public Economics*, 87(7–8), 1801–1824. https://doi.org/10.1016/S0047-2727(01)00186-4
- Bryant, Jennings, and Mary Beth Oliver, eds. Media effects: Advances in theory and research. Routledge, 2009.
- Budak, C., Nyhan, B., Rothschild, D. M.,

- Thorson, E., & Watts, D. J. (2024). Misunderstanding the harms of online misinformation. Nature, 630(8015), 45-53.
- Bullock, J. G. (2011). Elite influence on public opinion in an informed electorate. *American Political Science Review*, 105(3), 496–515.
- Bullock, J. G. (2020). Party cues. Edited by Suhay, Elizabeth, Bernard Grofman, and Alexander H. Trechsel. The Oxford Handbook of Electoral Persuasion (129-150). Oxford University Press.
- Bouwer, A. (2022). Under which conditions are humans motivated to delegate tasks to AI? A taxonomy on the human emotional state driving the motivation for AI delegation. In J. L. Reis, E. P. López, L. Moutinho, & J. P. M. d. Santos (Eds.), Marketing and smart technologies. smart innovation, systems and technologies (Vol. 279, 37–53). Springer, Singapore. https://doi.org/10.1007/978-981-16-9268-0_4
- Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. (1960). *The American voter*. John Wiley.
- Carpenter, P. (2024, October 2). *The liar's dividend: How AI is reshaping truth in business communications*. Forbes. https://www.forbes.com/councils/forbesbusinesscouncil/2024/10/02/the-liars-dividend-how-ai-isreshaping-truth-in-business-communications/
- Center for News, Technology & Innovation. (2024, October 11). *Synthetic media & deepfakes*. https://innovating.news/article/synthetic-media-deepfakes/
- Center for News, Technology & Innovation. (2024). *Addressing disinformation*. https://innovating.news/article/addressing-disinformation/
- Chen, Z., Kalla, J., Le, Q., Nakamura-Sakai, S., Sekhon, J., & Wang, R. (2025). *A framework to assess the persuasion risks large language model chatbots pose to democratic societies*.

- arXiv. https://arxiv.org/abs/2505.00036
- Chow, A. (2024). *AI's underwhelming impact* on the 2024 elections. Time. https://time.com/7131271/ai-2024-elections/
- Chowdhury, R. (2024). AI-fuelled election campaigns are here where are the rules? Nature, 28(8007), 237. https://doi.org/10.1038/d41586-024-00995-9
- Chu, X., Otto, L., Vliegenthart, R., Lecheler, S., de Vreese, C., & Kruikemeier, S. (2023). On or off topic? Understanding the effects of issue-related political targeted ads. *Information, Communication & Society, 27*(7), 1378–1404. https://doi.org/10.1080/13691 18X.2023.2265978
- Chung, M., Kim, N., Jones-Jang, S. M., Choi, J., & Lee, S. (2025). I see a double-edged sword: How self-other perceptual gaps predict public attitudes toward ChatGPT regulations and literacy interventions. *New Media & Society*, o(o). https://doi.org/10.1177/14614448241313180
- CISA. (2024). *Risk in focus: Generative A.I. and the 2024 election cycle*. https://www.cisa.gov/resources-tools/resources/risk-focusgenerative-ai-and-2024-election-cycle
- Computers are social actors. (2025, May 23). In *Wikipedia*. Retrieved May 30, 2025, from https://en.wikipedia.org/w/index.php?title=Computers_are_social_actors&oldid=1291858439
- Coppock, A. (2023). *Persuasion in parallel: How information changes minds about politics*. University of Chicago Press.
- Coppock, A., Green, D. P., & Porter, E. (2022). Does digital advertising affect vote choice? Evidence from a randomized field experiment. *Research & Politics*, *9*(1), 20531680221076901.
- Coppock, A., Hill, S. J., & Vavreck, L. (2020). The small effects of political advertising are small regardless of context, message,

- sender, or receiver: Evidence from 59 realtime randomized experiments. *Science Advances*, 6(36), eabc4046.
- Costello, T. H., Pennycook, G., & Rand, D. G. (2024). Durably reducing conspiracy beliefs through dialogues with AI. *Science*, 385(6714), eadq1814.
- Costello, T. H., Pennycook, G., Willer, R., & Rand, D. G. (2025). *Deep canvassing using AI*. OSF Preprints. https://osf.io/preprints/osf/q7e6u_v1
- Costello, T. H., Pennycook, G., & Rand, D. (2025). Just the facts: How dialogues with AI reduce conspiracy beliefs. *PsyArXiv*. https://osf.io/preprints/psyarxiv/h7n8u_v1
- Czarnek, G., Orchinik, R., Lin, H., Xu, H. G., Costello, T., Pennycook, G., & Rand, D. G. (2025). Addressing climate change skepticism and inaction using human-AI dialogues. *PsyArXiv*. https://osf.io/preprints/psyarxiv/mqcwj_v1
- Dahl, R. A. (1989). *Democracy and its critics*. Yale University Press.
- Davison, W. P. (1983). The third-person effect in communication. *Public Opinion Quarterly*, 47(1), 1–15.
- DellaVigna, S., & Gentzkow, M. (2010). Persuasion: Empirical evidence. *Annual Review of Economics*, *2*(1), 643–669.
- De Nadal, L., & Jančárik, P. (2024). *Beyond the deepfake hype: AI, democracy, and "the Slovak case."* Harvard Kennedy School (HKS) Misinformation Review. https://doi.org/10.37016/mr-2020-153
- Devine, C., O'Sullivan, D., & Lyngaas, S. (2024). A Fake Recording of a Candidate Saying He'd Rigged the Election Went Viral. Experts Say it's Only the Beginning. *CNN. https://edition.cnn.com/*2024/02/01/politics/election-deepfake-threats-invs
- Dobber, T., Ó Fathaigh, R., & Zuiderveen

- Borgesius, F. J. (2019). The regulation of online political micro-targeting in Europe. *Internet Policy Review*, 8(4). https://doi.org/10.14763/2019.4.1440
- Dommett, K., & Power, S. (2024). The shock of the old? The value of looking back when studying the mercurial world of political campaigning. *Political Communication Report*, 29(Spring), 1–8. https://doi.org/10.17169/refubium-43527
- Dommett, K., Kefford, G., & Kruschinski, S. (2024). *Data-driven campaigning and political parties: Five advanced democracies compared*. Oxford University Press.
- Drolsbach, C., & Pröllochs, N. (2025). *Characterizing AI-generated misinformation on social media*. arXiv. https://www.arxiv.org/abs/2505.10266
- Dubois, E., & Blank, G. (2018). The echo chamber is overstated: The moderating effect of political interest and diverse media. *Information, Communication & Society, 21*(5), 729–745. https://doi.org/10.1080/13691 18X.2018.1428656
- Dwoskin, E. (2024, January 22). *AI*deepfakes pose new challenges
 for politicians and elections. The
 Washington Post. https://www.washingtonpost.com/technology/2024/01/22/
 ai-deepfake-elections-politicians/
- Earp, B. D., Mann, S. P., Aboy, M., Awad, E., Betzler, M., Botes, M., Calcott, R., Caraccio, M., Chater, N., Coeckelbergh, M., Constantinescu, M., Dabbagh, H., Devlin, K., Ding, X., Dranseika, V., Everett, J. A. C., Fan, R., Feroz, F., Francis, ... & Clark, M. S. (2025). *Relational norms for human-AI cooperation*. arXiv. https://arxiv.org/abs/2502.12102
- Easton, M. J., & Holbein, J. B. (2021). The democracy of dating: How political affiliations shape relationship formation. *Journal of Experimental Political Science*, 8(3), 260–272. https://doi.org/10.1017/XPS.2020.21

- Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Kendeou, P., Vraga, E. K., & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, *1*(1), 13–29. https://doi.org/10.1038/s44159-021-00006-y
- Ejaz, W., Fletcher, R., Nielsen, R. K., & McGregor, S.(2024). What do people want? Views on platforms and the digital public sphere in eight countries. Reuters Institute for the Study of Journalism. https://doi.org/10.60625/risj-8pk9-d398
- Epstein, Z., Lin, H., Pennycook, G., & Rand, D. (2022). Quantifying attention via dwell time and engagement in a social media browsing environment. arXiv. https://arxiv.org/abs/2209.10464
- Elklit, J., & Reynolds, A. (2005). Judging elections and election management quality by process. *Representation*, *41*(3), 189–207.
- Falguera, E., Jones, S., & Ohman, M. (2014). Funding of political parties and election campaigns: A handbook on political finance. International IDEA.
- Fawzi, N., Steindl, N., Obermaier, M., Prochazka, F., Arlt, D., Blöbaum, B., Dohle, M., Engelke, K. M., Hanitzsch, T., Jackob, N., Jakobs, I., Klawier, T., Post, S., Reinemann, C., Schweiger, W., & Ziegele, M. (2021). Concepts, causes and consequences of trust in news media a literature review and framework. *Annals of the International Communication Association*, 45(2), 154–174. https://doi.org/10.1080/23808985.2021.1960181
- Federal Election Commission. (2025, January 28). *Statistical summary of 21-month campaign activity of the 2023-2024 election cycle*. https://www.fec.gov/updates/statistical-summary-of-21-month-campaign-activity-of-the-2023-2024-election-cycle/
- Fish, M., Wittenberg, J., & Jakli, L. (2018). *A decade of democratic decline and stagnation*. In C. Haerpfer, P. Bernhagen, R.

- Inglehart, & C. Welzel (Eds.), *Democratization* (2nd ed., pp. 268–284). Oxford University Press.
- Fletcher, R., & Nielsen, R. K. (2024). What does the public in six countries think of generative AI in news? Reuters Institute for the Study of Journalism. https://doi.org/10.60625/RISJ-4ZB8-CG87
- Fletcher, R., Andı, S., Badrinathan, S., Eddy, K. A., Kalogeropoulos, A., Mont'Alverne, C., Robertson, C. T., Arguedas, A. R., Schulz, A., Toff, B., & Nielsen, R. K. (2025). The link between changing news use and trust: Longitudinal analysis of 46 countries. *Journal of Communication*, 75(1), 1–15. https://doi.org/10.1093/joc/jqae044
- France24. (2025, February 12). *Scammers using AI to dupe the lonely looking for love*. https://www.france24.com/en/live-news/20250212-scammers-using-ai-to-dupe-the-lonely-looking-for-love
- Freille, S., Haque, M. E., & Kneller, R. (2007). A contribution to the empirics of press freedom and corruption. *European Journal of Political Economy*, 23(4), 838–862. https://doi.org/10.1016/j.ejpoleco.2007.03.002
- Fried, I. (2023, July 10). *How AI will tur-bocharge misinformation—And what we can do about it.* Axios. https://www.axios.com/2023/07/10/ai-misinformation-response-measures
- Garimella, K., & Chauchard, S. (2024). How prevalent is AI misinformation? What our studies in India show so far. *Nature*, 630(8015), 32–34. https://www.nature.com/articles/d41586-024-01588-2
- Gahn, C. (2024). How much tailoring is too much? Voter backlash on highly tailored campaign messages. The International Journal of Press/Politics. Advance online publication. https://doi.org/10.1177/19401612241263192
- Gilardi, F., Gessler, T., Kubli, M., & Müller, S.

- (2022). Social media and political agenda setting. *Political Communication*, 39(1), 39-60. https://doi.org/10.1080/10584609.2021 .1910390
- Gilardi, F., Kasirzadeh, A., Bernstein, A., & Vetere, F. (2024). We need to understand the effect of narratives about generative AI. *Nature Human Behaviour*, 8(12), 2251–2252. https://doi.org/10.1038/s41562-024-02026-z
- Ginsburg, T., & Huq, A. (2018). *How to save a constitutional democracy*. The University of Chicago Press.
- Gold, A. & Fischer, S. (2023, February 21). *Chatbots trigger next misinformation nightmare*. Axios. https://www.axios.com/2023/02/21/chatbots-misinformation-nightmare-chatgpt-ai
- Goldman, S., & Kahn, J. (2025, April 16). *OpenAI* updated its safety framework but no longer sees mass manipulation and disinformation as a critical risk. Fortune. https://fortune.com/2025/04/16/openai-safety-framework-manipulation-deception-critical-risk/
- Goldstein, J. A., & Lohn, A. (2024, January 23). *Deepfakes, elections, and shrinking the liar's dividend*. Brennan Center for Justice. https://www.brennancenter.org/our-work/research-reports/deepfakes-elections-and-shrinking-liars-dividend
- Grace, K., Stewart, H., Sandkühler, J. F., Thomas, S., Weinstein-Raun, B., & Brauner, J. (2024). *Thousands of AI authors on the future of AI*. arXiv. https://arxiv.org/abs/2401.02843
- Green, D., Palmquist, B., & Schickler, E. (2002). Partisan hearts and minds: Political parties and the social identities of voters. Yale University Press.
- Greenwood, M. (2025, May 23). *Most political consultants are using AI, study finds*. Campaigns & Elections. https://campaignsandelections.com/industry-news/

81

- study-finds-most-political-consultants-using-ai/
- Guess, A. M., Nyhan, B., O'Keeffe, Z., & Reifler, J. (2020). The sources and correlates of exposure to vaccine-related (mis) information online. Vaccine, 38(49), 7799-7805.
- Habgood-Coote, J. (2023). Deepfakes and the epistemic apocalypse. *Synthese*, 201(103). https://doi.org/10.1007/s11229-023-04097-3
- Hackenburg, K., & Margetts, H. (2024a).
 Evaluating the persuasive influence of political microtargeting with large language models. *Proceedings of the National Academy of Sciences of the United States of America*, 121(24), e2403116121. https://doi.org/10.1073/pnas.2403116121
- Hackenburg, K., & Margetts, H. (2024b). Reply to Teeny and Matz: Toward the robust measurement of personalized persuasion with generative AI. *Proceedings of the National Academy of Sciences of the United States of America*, 121(43), e2418817121. https://doi.org/10.1073/pnas.2418817121
- Hackenburg, K., Tappin, B., Röttger, P., Hale, S., Bright, J., & Margetts, H. (2025). *Scaling language model size yields diminishing returns for single message political persuasion*. Proceedings of the National Academy of Sciences of the United States of America. https://eprints.lse.ac.uk/127194/
- Haenschen, K. (2023). The conditional effects of microtargeted Facebook advertisements on voter turnout. *Political Behavior*, *45*, 1661–1681. https://doi.org/10.1007/s11109-022-09781-7
- Hager, A. (2019). Do online ads influence vote choice? *Political Communication*, *36*(3), 376–393.
- Hansen, K. M., & Pedersen, R. T. (2014). Campaigns matter: How voters become knowledgeable and efficacious during election campaigns. *Political Communication*, *31*(2), 303–324. https://doi.org/10.1080/10584609.

2013.815296

- Hameleers, M., van der Meer, T. G. L. A., & Dobber, T. (2023). They would never say anything like this! Reasons to doubt political deepfakes. *European Journal of Communication*, *39*(1), 56–70. https://doi.org/10.1177/02673231231184703
- Harris, P. L. (2012). *Trusting what you're told: How children learn from others*. Harvard University Press.
- Harris, K.R. (2021). Video on demand: what deepfakes do and how they harm. *Synthese*, 199, 13373–13391. https://doi.org/10.1007/S11229-021-03379-y
- Hartmann, D., Pohlmann, L., Wang, S. M., & Berendt, B. (2024). A systematic review of echo chamber research: Comparative analysis of conceptualizations, operationalizations, and varying outcomes. arXiv. https://arxiv.org/abs/2407.06631
- Hennemann-Heldt, A. (2024). *Understanding the role of generative AI in elections: A crucial endeavor in 2024*. TUM Think Tank. https://tumthinktank.de/wp-content/uploads/GenAIElections Report TTT.pdf
- Herre, B., Rodés-Guirao, L., & Ortiz-Ospina, E. (2024). *Democracy*. Our World in Data. https://ourworldindata.org/democracy?insight=the-world-has-recently-become-less-democratic#key-insights
- Hersh, E. D. (2015). *Hacking the electorate: How campaigns perceive voters*. Cambridge University Press.
- Hersh, E. D., & Schaffner, B. F. (2013). Targeted campaign appeals and the value of ambiguity. *Journal of Politics*, 75(2), 520–534. https://doi.org/10.1017/S0022381613000182
- Hill, K. (2025, June 13). They asked an A.I. chatbot questions. The answers sent them spiraling. *The New York Times*. https://www.nytimes.com/2025/06/13/technology/chatgpt-ai-chatbots-conspiracies.html

- HGL Team. (2024, May 23). *AI Edition: Higher ground labs' political tech landscape report*. Higher Ground Labs. https://higherground-labs.com/ai-landscape-report/
- Hoffmann, C. P. (2023). The role of moral panics in media transformation: An examination of the "techlash." In A. Godulla & S. Böhm (Eds.), *Digital disruption and media transformation: How technological innovation shapes the future of communication* (pp. 41–54). Springer International Publishing.
- House of Lords Communications and Digital Committee: Large language models and generative AI (1st Report of Session 2023–24; House of Lords Communications and Digital Committee). (2024). House of Lords. https://publications.parliament.uk/pa/ld5804/ldselect/ldcomm/54/5402.htm
- Hsu, T., & Thompson, S. A. (2023, February 8).
 AI chatbots could spread disinformation,
 experts warn. The New York Times. https://
 www.nytimes.com/2023/02/08/technology/
 ai-chatbotsdisinformation.html
- Huang, H., & Cruz, N. (2022). Propaganda, presumed influence, and collective protest. *Political Behavior*, *44*(4), 1789–1812. https://doi.org/10.1007/s11109-021-09683-0
- Huber, G. A., & Malhotra, N. (2017). Political homophily in social relationships: Evidence from online dating behavior. *Journal of Politics*, 79(1), 269–283. https://doi.org/10.1086/687533
- Hudde, A., & Grunow, D. (2025). Why do partners often prefer the same political parties? Evidence from couples in Germany. *Social Forces*, 103(4), 1581–1603. https://doi.org/10.1093/sf/soae133
- Huddy, L., Sears, D. O., Levy, J. S., & Jerit, J. (Eds.). (2023). *The Oxford Handbook of Political Psychology* (3rd ed.). Oxford University Press.
- Hughes, T. P. (1994). Technological momentum. In L. Marx & M. Roe Smith (Eds.), *Does*

- technology drive history? The dilemma of technological determinism (pp. 101–113). MIT Press.
- Hughes, T. P. (2012). The evolution of large technological systems. In W. E. Bijker, T. P. Hughes, & T. Pinch (Eds.), *The social construction of technological systems* (pp. 45–76). MIT Press.
- Humprecht, E., Esser, F., & Van Aelst, P. (2020). Resilience to online disinformation: A framework for cross-national comparative research. *The International Journal of Press/Politics*, 25(3), 493-516.
- Hwang, T. (2020). Subprime attention crisis: Advertising and the time bomb at the heart of the Internet. FSG Originals.
- Ibrahim, L., Akbulut, C., Elasmar, R., Rastogi, C., Kahng, M., Morris, M. R., McKee, K. R., Rieser, V., Shanahan, M., & Weidinger, L. (2025). *Multi-turn evaluation of anthropomorphic behaviours in large language models*. arXiv. https://arxiv.org/abs/2502.07077
- Information Commissioner's Office. (n.d.). *Microtargeting*. https://ico.org.uk/for-the-public/microtargeting/
- Jaźwińska, K., & Chandrasekar, A. (2025, March 6). AI search has a citation problem. Columbia Journalism Review. https://www.cjr.org/tow_center/we-compared-eight-ai-search-engines-theyre-all-bad-at-citing-news.php
- Jennings, M. K, & Stoker, L. (2001). *Political* similarity and influence between husbands and wives. UC Berkeley: Institute of Governmental Studies. https://escholarship.org/uc/item/9s54f2mc
- Jones, C. R., & Bergen, B. K. (2024). Lies, damned lies, and distributional language statistics: Persuasion and deception with large language models. arXiv. https://arxiv.org/abs/2412.17128
- Jungherr, A. (2023). Artificial intelligence and democracy: A conceptual framework. *Social*

- *Media + Society*, *9*(3), 20563051231186353. https://doi.org/10.1177/20563051231186353
- Jungherr, A., & Rauchfleisch, A. (2024). Negative downstream effects of alarmist disinformation discourse: Evidence from the United States. *Political Behavior*, 46:2123–2143.
- Jungherr, A., Rivero, G., & Gayo-Avello, D. (2020). *Retooling politics: How digital media are shaping democracy*. Cambridge University Press.
- Jungherr, A., & Schroeder, R. (2021). Disinformation and the structural transformations of the public arena: Addressing the actual challenges to democracy. Social Media+Society, 7(1), 2056305121988928.
- Kalla, J. L., & Broockman, D. E. (2018). The minimal persuasive effects of campaign contact in general elections: Evidence from 49 field experiments. *American Political Science Review*, 112(1), 148–166.
- Kapoor, S., & Narayanan, A. (2024, December 13). We looked at 78 election deepfakes. Political misinformation is not an AI problem. AI Snake Oil. https://www.aisnakeoil.com/p/we-looked-at-78-election-deepfakes
- Narayanan, A., & Kapoor, S. (2024). AI Snake Oil: What Artificial Intelligence Can Do, What It Can't, and How to Tell the Difference. Princeton University Press.
- Kahloon, I., & Ramani, A. (2023, August 31). *AI* will change American elections, but not in the obvious way. The Economist. https://www.economist.com/united-states/2023/08/31/ai-will-change-american-elections-but-not-in-the-obvious-way
- Karpf, D. (2019, December 10). On digital disinformation and democratic myths. MediaWell, Social Science Research Council. https://mediawell.ssrc.org/expert-reflections/on-digital-disinformation-and-democratic-myths/
- Kefford, G., Dommett, K., Baldwin-Philippi, J.,

- Bannerman, S., Dobber, T., Kruschinski, S., Kruikemeier, S., & Rzepecki, E. (2022). Datadriven campaigning and democratic disruption: Evidence from six advanced democracies. *Party Politics*, *29*(3), 448-462. https://doi.org/10.1177/13540688221084039
- Kelley, J. (2012). Monitoring democracy: When international election observation works, and why it often fails. Princeton University Press.
- Kirk, H. R., Gabriel, I., Summerfield, C., Vidgen, B., & Hale, S. A. (2025). *Why human-AI relationships need socioaffective alignment*. arXiv. https://arxiv.org/abs/2502.02528
- Kreiss, D., & McGregor, S. C. (2018). Technology firms shape political communication: The work of Microsoft, Facebook, Twitter, and Google with campaigns during the 2016 US presidential cycle. *Political Communication*, 35(2), 155–177. https://doi.org/10.1080/10584609.2017.1364814
- Kreiss, D. (2016). *Prototype politics: Technology intensive campaigning and the data of democracy*. Oxford University Press.
- Knight, T. (2024). South Africa's election marked by an unexpected lack of artificially generated content. In I. K. Trauthig & S. Woolley (Eds.), Series on generative artificial intelligence and elections. Center for Media Engagement. https://mediaengagement. org/research/generative-artificial-intelligence-and-elections
- Knight, W. (2016, October 26). *Chatbots with social skills will convince you to buy something*. MIT Technology Review. https://www.technologyreview.com/2016/10/26/156438/chatbots-with-social-skills-will-convince-you-to-buy-something/
- Kruschinski, S., & Haller, A. (2017). Data driven online political micro targeting: Hunting for voters, shooting democracy? Centre for Media Pluralism and Media Freedom Report. https://cmpf.eui.eu/data-driv-en-online-political-microtargeting-hunt-

- ing-for-voters-shooting-democracy/
- Lambrecht, A., & Tucker, C. (2019). Algorithmic bias? An empirical study of apparent gender-based discrimination in the display of STEM career ads. *Management Science*, 65(7), 2966–2981. https://doi.org/10.1287/mnsc.2018.3093
- Leake, M. (2024). Are fears about online misinformation in the US election overblown? The evidence suggests they might be. Reuters Institute. https://reutersinstitute.politics. ox.ac.uk/news/are-fears-about-online-misinformation-us-election-overblown-evidence-suggests-they-might-be
- Leighley, J. E., & Nagler, J. (2014). Who votes now?: Demographics, issues, inequality, and turnout in the United States. Princeton University Press.
- Leonardi P. M. (2012). Car crashes without cars: Lessons about simulation technology and organizational change from automotive design. MIT Press.
- Levitsky, S., & Way, L. (2010). *Competitive authoritarianism: Hybrid regimes after the Cold War*. Cambridge University Press.
- Levitsky, S., & Ziblatt, D. (2018). *How democracies die*. Crown.
- Lorenz-Spreen, P., Mørch Mønsted, B., Hövel, P., & Lehmann, S. (2019). Accelerating dynamics of collective attention. *Nature Communications*, 10(1759). https://doi.org/10.1038/s41467-019-09311-w
- Marcus, G. (2023, February 8). AI's Jurassic Park moment. *Communications of the ACM*. https://cacm.acm.org/blogs/blogcacm/267674-ais-jurassic-park-moment/fulltext
- Marinov, V. (2024, May 23). *Don't bother* asking AI about the EU elections: How chatbots fail when it comes to politics. Correctiv. https://correctiv.org/en/fact-checking-en/2024/05/23/

- dont-bother-askingai-about-the-eu-elections-how-chatbots-fail-whenit-comes-to-politics/
- McClain, C. (2024). *Americans' use of ChatGPT is ticking up, but few trust its election information*. Pew Research Center. https://www.pewresearch.org/short-reads/2024/03/26/americans-use-of-chatgpt-is-ticking-up-but-few-trust-its-election-information/
- Mansbridge, J. (1999). Should blacks represent blacks and women represent women? A contingent "yes." *Journal of Politics*, *61*(3), 628–657.
- Mazepus, H., Osmudsen, M., Bang-Petersen, M., Toshkov, D., & Dimitrova, A. (2023). Information battleground: Conflict perceptions motivate the belief in and sharing of misinformation about the adversary. *PLOS ONE*, *18*(3), e0282308. https://doi.org/10.1371/journal.pone.0282308
- Media Cloud. (n.d.). *Global English Language Sources database*. https://search.mediacloud.
 org/
- Mimizuka, K., Brown, M. A., Yang, K-C., & Lukito, J. (2025). *Post-post-API age: Studying digital platforms in scant data access times*. arXiv. https://arxiv.org/abs/2505.09877
- Mercier, H. (2020). *Not born yesterday: The science of who we trust and what we believe.* Princeton University Press.
- Mercier, H., & Sperber, D. (2017). *The enigma of reason*. Harvard University Press.
- Mercier, H., & Claidière, N. (2022). Does discussion make crowds any wiser? *Cognition*, 222, 104912.
- Meta (2024, December 3) What we saw on our platforms during 2024's global elections. Meta. https://about.fb.com/news/2024/12/2024-global-elections-metaplatforms/
- Mitchell, M. (2019). Artificial Intelligence: A

- Guide for Thinking Humans. Pelican.
- Morin, O., Jacquet, P. O., Vaesen, K., & Acerbi, A. (2021). Social information use and social information waste. *Philosophical Transactions of the Royal Society B*, 376(1828), 20200052.
- Motta, M., Hwang, J., & Stecula, D. (2024). What goes down must come up? Pandemic-related misinformation search behavior during an unplanned Facebook outage. *Health Communication*, 39(10), 2041–2052.
- Mummolo, J., Peterson, E., & Westwood, S. (2021). The limits of partisan loyalty. *Political Behavior*, *43*(3), 949–972.
- Murgia, M. (2024). *Code dependent: Living in the shadow of AI*. Picador.
- Nicas, J., & Herrera, L. C. (2023). *Is Argentina the first AI election?* The New York Times. https://www.nytimes.com/2023/11/15/world/americas/argentina-election-ai-milei-massa.html
- Nielsen, R. (2024, January 2). Forget technology politicians pose the gravest misinformation threat. Financial Times. https://www.ft.com/content/5da52770-b474-4547-8d1b-9c46a3c3bac9
- Nicholson, S. P., Coe, C. M., Emory, J., & Song, A. V. (2016). The politics of beauty: The effects of partisan bias on physical attractiveness. *Political Behavior*, *38*(4), 883–898. https://doi.org/10.1007/s11109-016-9339-7
- Newman, N., Fletcher, R., Robertson, C. T., Ross Arguedas, A., & Nielsen, R. K. (2024). *Digital News Report 2024*. Reuters Institute for the Study of Journalism.
- Newman, N., Fletcher, R., Robertson, C. T., Ross Arguedas, A., & Nielsen, R. K. (2025). *Digital News Report 2025*. Reuters Institute for the Study of Journalism.
- News Literacy Project (2024, October 29)

 Experts: Watch out for AI-generated fakes

- and disinformation about voting ahead of election day. https://newslit.org/newsroom/press-release/the-news-liter-acy-project-experts-watch-out-for-ai-generated-fakes-and-disinformation-about-voting-ahead-of-election-day/
- Norris, P. (2015). *Why elections fail*. Cambridge University Press.
- Nyhan, B., Porter, E., Reifler, J., & Wood, T. J. (2020). Taking fact-checks literally but not seriously? The effects of journalistic fact-checking on factual beliefs and candidate favorability. *Political Behavior*, 42, 939–960. https://doi.org/10.1007/s11109-019-09528-x
- Nyhan, B., Settle, J., Thorson, E., Wojcieszak, M., Barberá, P., Chen, A. Y., Allcott, H.,
- Brown, T., Crespo-Tenorio, A., Dimmery, D., Freelon, D., Gentzkow, M., González-Bailón, S., Guess, A. M., Kennedy, E., Kim, Y. M., Lazer, D., Malhotra, N., Moehler, D., ... & Tucker, J. A. (2023). Like-minded sources on Facebook are prevalent but not polarizing. *Nature*, 620(7972), 137–144.
- Office of the Director of National Intelligence (ODNI). (2024, September). *Election security update as of mid-September 2024* (OV-2024-26092). https://www.dni.gov/files/FMIC/documents/ODNI-Election-Security-Update-20240923.pdf
- *OECD AI Principles*. (2024). https://oecd.ai/en/principles
- Orben, A. (2020). The Sisyphean cycle of technology panics. *Perspectives on Psychological Science*, *15*(5), 1143–1157.
- Orchinik, R., Bhui, R., & Rand, D. G. (2024). *Repetition does not increase belief in claims from distrusted politicians*. PsyArXiv. https://osf.io/preprints/psyarxiv/y7pn5_v1
- Orchinik, R., Rand, D., & Bhui, R. (2024). The not so illusory truth effect: A rational foundation for repetition effects. PsyArXiv.

86

- https://osf.io/preprints/psyarxiv/qvwdy_v3
- Ordonez, V., Dunn, T., & Noll, E. (2023, May 19). OpenAI CEO Sam Altman says AI will reshape society, acknowledges risks: 'A little bit scared of this'. ABC News. https://abcnews.go.com/Technology/openai-ceo-sam-altman-ai-reshape-societyacknowledges/story?id=97897122
- Oser, J., Grinson, A., Boulianne, S., & Halperin, E. (2022). How political efficacy relates to online and offline political participation: A multilevel meta-analysis. *Political Communication*, 39(5), 607–633. https://doi.org/10.1080/10584609.2022.2086329
- Parasie S., Machut A., & Mazoyer B. (2025). *The politicisation of news stories on social networks*. SciencesPo Médialab. https://www.sciencespo.fr/recherche/sites/sciencespo.fr.recherche/files/CST2_The%20politicisationpdf.pdf#page=8.00
- Paris, B., & Donovan, J. (2019). *Deepfakes and cheapfakes: The manipulation of audio and visual evidence*. Data & Society Research Institute. https://datasociety.net/library/deepfakes-and-cheap-fakes/
- Pasternack, A. (2023, March 17). *Deepfakes getting smarter thanks to GPT*. FastCompany. https://www.fastcompany.com/90853542/deepfakes-getting-smarter-thanks-to-gpt
- Peresman, A., Larsen, L. T., Mazepus, H., & Petersen, M. B. (2025). Do populists listen to expertise? A five-country study of authority, arguments, and expert sources. *Political Studies*, *o*(o). https://doi.org/10.1177/00323217251323402
- Petersen, M. B., Osmundsen, M., & Arceneaux, K. (2023). The "need for chaos" and motivations to share hostile political rumors. *American Political Science Review*, 117(4), 1486–1505. https://doi.org/10.1017/S0003055422001447
- Petre, C. (2021). All the news that's fit to click:

- How metrics are transforming the work of journalists. Princeton University Press.
- Pfänder, J., & Altay, S. (2025). Spotting false news and doubting true news: a systematic review and meta-analysis of news judgements. *Nature human behaviour*, *9*(4), 688–699. https://doi.org/10.1038/s41562-024-02086-1
- Pillai, R. M., Fazio, L. K., & Effron, D. A. (2023). Repeatedly encountered descriptions of wrongdoing seem more true but less unethical: Evidence in a naturalistic setting. *Psychological Science*, *34*(8), 863–874.
- Pillai, R. M., Kim, E., & Fazio, L. K. (2023). All the president's lies: Repeated false claims and public opinion. *Public Opinion Quarterly*, 87(3), 764–802.
- Porter, E., & Wood, T. J. (2024). Factual corrections: Concerns and current evidence. *Current Opinion in Psychology*, *55*, 101715.
- Radcliffe, D. (2025). *Journalism in the AI Era: Opportunities and Challenges in the Global South and Emerging Economies*. Thomson
 Reuters Foundation.
- Rahman-Jones, I. (2025, February 11). *AI chat-bots unable to accurately summarise news, BBC finds.* BBC News. https://www.bbc.co.uk/news/articles/com17d8827ko
- Rahman, S., Dip, S., Karim, F., & Naser, S. (2024). Whisperers amidst the echo chamber: Decoding Bangladesh's 2024 election disinformation. Center for Critical and Qualitative Studies, University of Liberal Arts Bangladesh. https://shorturl.at/adptS
- Rain, M., & Mar, R. A. (2021). Adult attachment and engagement with fictional characters. *Journal of Social and Personal Relationships*, 38(9), 2792–2813.
- Rrv, A., Tyagi, N., Uddin, M. N., Varshney, N., & Baral, C. (2024). Chaos with keywords: Exposing large language models' sycophancy to misleading keywords and

- evaluating defense strategies. In *Findings of the Association for Computational Linguistics: ACL 2024* (pp. 12717–12733). Association for Computational Linguistics. https://aclanthology.org/2024.findings-acl.755/
- Robertson, R. E., Green, J., Ruck, D. J., Ognyanova, K., Wilson, C., & Lazer, D. (2023). Users choose to engage with more partisan news than they are exposed to on Google search. *Nature*, *618*(7964), 342–348.
- Ross Arguedas, A., Robertson, C. T., Fletcher, R., & Nielsen, R. K. (2022, January 19). *Echo chambers, filter bubbles, and polarisation: A literature review*. Reuters Institute for the Study of Journalism. https://doi.org/10.60625/risj-etxj-7k60
- Ryazanov, I., Öhman, C., & Björklund, J. (2025). How ChatGPT changed the media's narratives on AI: A semi automated narrative analysis through frame semantics. *Minds & Machines*, 35(2). https://doi.org/10.1007/S11023-024-09705-W
- Russell, S. J., & Norvig, P. (2009). *Artificial Intelligence: A Modern Approach* (3rd ed.). Pearson.
- Safiullah, M., & Parveen, N. (2022). *Big data, artificial intelligence and machine learning: A paradigm shift in election campaigns*. In S. K. Panda, R. K. Mohapatra, S. Panda, & S. Balamurugan (Eds.), *The new advanced society: Artificial intelligence and industrial internet of things paradigm* (pp. 247–262). Scrivener Publishing. https://doi.org/10.1002/9781119884392.ch11
- Sehgal, N. K. R., Rai, S., Tonneau, M., Agarwal, A. K., Cappella, J., Kornides, M., Ungar, L., Buttenheim, A., Guntuku, S. C. (2025) Conversations with AI chatbots increase short-term vaccine intentions but do not outperform standard public health messaging. ArXiv. https://arxiv.org/abs/2504.20519
- Shanahan, M., McDonell, K., & Reynolds, L. (2023, November 8). Role play with large language models. *Nature*, 623, 493–498.

- https://doi.org/10.1038/s41586-023-06647-8
- Schiff, K. J., Schiff, D. S., & Bueno, N. S. (2025). The liar's dividend: Can politicians claim misinformation to evade accountability? *American Political Science Review*, 119(1), 71–90. https://doi.org/10.1017/S0003055423001454
- Schroeder, R. (2018). Social theory after the Internet: Media, technology, and globalization. UCL Press.
- Scott, L. (2023, September 5). World faces "tech enabled Armageddon," Maria Ressa says.

 Voice of America. https://www.voanews.
 com/a/world-faces-tech-enabled-armageddon-maria-ressa-says-/7256196.html
- Sethuraman, R., Tellis, G. J., & Briesch, R. A. (2011). How well does advertising work? Generalizations from meta analysis of brand advertising elasticities. Journal of Marketing Research, 48(3), 457–471.
- Shah, C., & Bender, E. (2023). *Envisioning* information access systems: What makes for good tools and a healthy web? Unpublished manuscript. https://faculty.washington.edu/ebender/papers/Envisioning_IAS_preprint.pdf
- Sharma, M., Tong, M., Korbak, T., Duvenaud, D. K., Askell, A., Bowman, S. R., Cheng, N., Durmus, E., Hatfield-Dodds, Z., Johnston, S., Kravec, S., Maxwell, T., McCandlish, S., Ndousse, K., Rausch, O., Schiefer, N., Yan, D., & Zhang, M. (2023). *Towards understanding sycophancy in language models*. arXiv. https://arxiv.org/abs/2310.13548
- Shapiro, B. T., Hitsch, G. J., & Tuchman, A. E. (2021). TV advertising effectiveness and profitability: Generalizable results from 288 brands. *Econometrica*, 89, 1855–1879. https://doi.org/10.3982/ECTA17674
- Simon, F. M. (2019). "We power democracy": Exploring the promises of the political data analytics industry. *The Information Society*, 53(3), 1–13.

- Simon, F. M. (2025). Rationalisation of the news: How AI reshapes and retools the gatekeeping processes of news organisations in the United Kingdom, United States and Germany. *New Media & Society*, 14614448251336423. https://doi.org/10.1177/14614448251336423
- Simon, F. M. (2024). *Artificial intelligence in the news. How AI retools, rationalizes, and reshapes journalism and the public arena* (p. 46). Tow Center for Digital Journalism, Columbia University. https://doi.org/10.7916/ncm5-3v06
- Simon, F. M., Adami, M., Kahn, G., & Fletcher, R. (2024). *How AI chatbots responded to basic questions about the 2024 European elections right before the vote*. Reuters Institute for the Study of Journalism. https://reutersinstitute.politics.ox.ac.uk/news/how-ai-chatbots-responded-basicquestions-about-2024-european-elections-right-vote
- Simon, F. M., Fletcher, R., & Nielsen, R. K. (2024). How generative AI chatbots responded to questions and fact checks about the 2024 UK general election. Reuters Institute for the Study of Journalism. https://reutersinstitute.politics.ox.ac.uk/how-generative-ai-chatbots-responded-questions-and-fact-checks-about-2024-uk-general-election
- Simon, F. M., & Camargo, C. Q. (2021). Autopsy of a metaphor: The origins, use and blind spots of the "infodemic." *New Media & Society*, *25*(8), 2219–2240. https://doi.org/10.1177/14614448211031908
- Simon, F. M., Altay, S., & Mercier, H. (2023).

 Misinformation reloaded? Fears about the impact of generative AI on misinformation are overblown. *Harvard Kennedy School Misinformation Review*, 4(5). https://misinforeview.hks.harvard.edu/article/misinformation-reloaded-fears-about-the-impact-of-generative-ai-on-misinformation-are-overblown/

- Sichova, A., & Das, P. (2024). *'AI armageddon':*Did AI misinformation really sway voters in 2024? Logically Facts. https://www.logicallyfacts.com/en/analysis/ai-armageddon-did-ai-misinformation-really-sway-voters-in-2024
- Simchon, A., Edwards, M., & Lewandowsky, S. (2024). The persuasive effects of political microtargeting in the age of generative artificial intelligence. PNAS nexus, 3(2), pgae035
- Snyder, J. M., Jr., & Strömberg, D. (2010). Press coverage and political accountability. *Journal of Political Economy*, *118*(2), 355–403. https://doi.org/10.1086/652903
- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language*, 25(4), 359-393. https://doi.org/10.1111/j.1468-0017.2010.01394.x
- Skjuve, M., Følstad, A., Fostervold, K. I., & Brandtzaeg, P. B. (2022). A longitudinal study of human–chatbot relationships. *International Journal of Human-Computer Studies*, 168, 102903. https://doi.org/10.1016/j.ijhcs.2022.102903
- Slothuus, R., & Bisgaard, M. (2021). Party over pocketbook? How party cues influence opinion when citizens have a stake in policy. *American Political Science Review*, 115(3), 1090–1096. https://doi.org/10.1017/S0003055421000332
- Spring, M. (2024). *This wasn't the social media election everyone expected*. BBC. https://www.bbc.com/news/articles/cj50qjy9g7ro
- Staab, R., Vero, M., Balunović, M., & Vechev, M. (2024). *Beyond memorization: Violating privacy via inference with large language models* [Conference paper]. Twelfth International Conference on Learning Representations (ICLR 2024). OpenReview. https://doi.org/10.3929/ethz-b-000720050
- Stockwell, S. (2024). AI-enabled influence

- operations: Threat analysis of the 2024 UK and European elections. Centre for Emerging Technology and Security. https://cetas.turing.ac.uk/publications/ai-enabled-influence-operations-threat-analysis-2024-uk-and-european-elections
- Shukla, V., & Schneier, B. (2024). Indian election was awash in deepfakes—but AI was a net positive for democracy. *The Conversation*. https://theconversation.com/indian-election-was-awash-in-deepfakes-but-ai-was-anet-positive-for-democracy-231795
- Strömbäck, J., Tsfati, Y., Boomgaarden, H., Damstra, A., Lindgren, E., Vliegenthart, R., & Lindholm, T. (2020). News media trust and its impact on media use: toward a framework for future research. *Annals of the International Communication Association*, 44(2), 139–156. https://doi.org/10.1080/23808985.2020.1755338
- Tappin, B. M., & McKay, R. (2021). Estimating the causal effects of cognitive effort and policy information on party cue influence. Psy-ArXiv. https://osf.io/preprints/psyarxiv/tdk3y_v1
- Tappin, B. M., Berinsky, A. J., & Rand, D. G. (2023). Partisans' receptivity to persuasive messaging is undiminished by countervailing party leader cues. *Nature Human Behaviour*, 7(4), 568–582.
- Taylor, G. (2014). Scarcity of Attention for a Medium of Abundance. An Economic Perspective. In M. Graham & W. H. Dutton (Eds.), *Society & The Internet* (pp. 257–271). Oxford University Press.
- Tenenboim-Weinblatt, K., Baden, C., Aharoni, T., & Overbeck, M. (2022). Affective forecasting in elections: A socio-communicative perspective. *Human Communication Research*, *48*(4), 553–566. https://academic.oup.com/hcr/article/48/4/553/657
- Thorson, E. (2024). *How news coverage of mis- information shapes perceptions and trust.* Cambridge University Press.

- Tiku, N. (2025, May 31). Your chatbot friend might be messing with your mind. The Washington Post. https://www.washingtonpost.com/technology/2025/05/31/ai-chatbots-user-influence-attention-chatgpt/
- Toff, B., & Simon, F. M. (2024). "Or they could just not use it?": The dilemma of AI disclosure for audience trust in news. The International Journal of Press/Politics. https://doi.org/10.1177/19401612241308697
- Törnberg, P., & Chueri, J. (2025). When do parties lie? Misinformation and radical-right populism across 26 countries. *The International Journal of Press/Politics*, *o*(o). https://doi.org/10.1177/194016122413
- Tucker, J. (2023, July 14). *AI could create a disinformation nightmare in the 2023 election*. The Hill. https://thehill.com/opinion/4096006-ai-could-create-a-disinformation-nightmare-in-the-2024election/
- Tuquero, L. (2024, December 25). Did artificial intelligence shape the 2024 US election? *Al Jazeera*. https://www.aljazeera.com/news/2024/12/25/did-artificial-intelligence-shape-the-2024-us-election
- Vaccari, C., & Valeriani, A. (2021). *Outside the bubble: Social media and political participation in Western democracies*. Oxford University Press.
- Valgarðsson, V., Jennings, W., Stoker, G., Bunting, H., Devine, D., McKay, L., & Klassen, A. (2025). A crisis of political trust? Global trends in institutional trust from 1958 to 2019. *British Journal of Political Science*, 55(3), 856-877. https://doi.org/10.1017/S0007123424000498
- Verma, P., & Zakrzewski, C. (2024, April 23). *AI deepfakes pose new challenges for politicians and elections*. The Washington Post. https://www.washingtonpost.com/technology/2024/04/23/ai-deepfake-election-2024-us-india/
- Vliegenthart, R., Vrielink, J., Dommett, K.,

- Gibson, R., Bon, E., Chu, X., de Vreese, C., Lecheler, S., Matthes, J., Minihold, S., Otto, L., Stubenvoll, M., & Kruikemeier, S. (2024). Citizens' acceptance of data-driven political campaigning: A 25-country cross-national vignette study. *Social Science Computer Review*, *42*(5), 1101–1119. https://doi.org/10.1177/08944393241249708
- Vogels, E. A. (2022, May 13). Support for more regulation of tech companies has declined in U.S., especially among Republicans. Pew Research Center. https://www.pewresearch.org/short-reads/2022/05/13/support-formore-regulation-of-tech-companies-has-declined-in-u-s-especially-among-republicans/
- Votta, F., Noroozian, A., Dobber, T., Helberger, N., & de Vreese, C. (2023). Going micro to go negative? Targeting toxicity using Facebook and Instagram ads. *Computational Communication Research*, 5(1), 1–50.
- Votta, F., Dobber, T., Guinaudeau, B., Helberger, N., & de Vreese, C. (2024). The cost of reach: Testing the role of ad delivery algorithms in online political campaigns. *Political Communication*, *42*(3), 476–508. https://doi.org/10.1080/10584609.2024.2439317
- Votta, F., Kruschinski, S., Hove, M., Helberger, N., Dobber, T., & de Vreese, C. (2024). Who does(n't) target you? Mapping the worldwide usage of online political microtargeting. *Journal of Quantitative Description:*Digital Media, 4. https://doi.org/10.51685/jqd.2024.010
- Wellman, B. (2004). The three ages of internet studies: Ten, five and zero years ago. *New Media & Society*, 6, 123–129.
- West, D. M., & Lo, J. (2024, January 23). *Misunderstood mechanics: How AI, TikTok, and the liar's dividend might affect the 2024 elections*. Brookings Institution. https://www.brookings.edu/articles/misunderstood-mechanics-how-ai-tiktok-and-the-liars-dividend-might-affect-the-2024-elections/

- Weikmann, T., Greber, H., & Nikolaou, A. (2024). After deception: How falling for a deepfake affects the way we see, hear, and experience media. *The International Journal of Press/Politics*, 30(1), 187–210. https://doi.org/10.1177/19401612241233539
- Williams, D. (2022). The focus on misinformation leads to a profound misunderstanding of why people believe and act on bad information. Impact of Social Sciences. London School of Economics and Political Science. https://blogs.lse.ac.uk/impactofsocialsciences/2022/09/05/thefocus-on-misinformation-leads-to-a-profound-misunderstanding-of-why-people-believe-and-act-on-bad-information/
- Williams, D. (2023). The marketplace of rationalizations. *Economics & Philosophy*, 39(1), 99–123.
- Williams, M., Carroll, M., Narang, A., Weisser, C., Murphy, B., & Dragan, A. (2025). *On targeted manipulation and deception when optimizing LLMs for user feedback*. arXiv. https://arxiv.org/abs/2411.02306
- Wong, M. (2024). Al's fingerprints were all over the election: But deepfakes and disinformation weren't the main issues. The Atlantic. https://www.theatlantic.com/technology/archive/2024/11/ai-election-propaganda/680677/
- World Economic Forum. (2024, January 10). *Global risks report 2024*. https://www.weforum.org/publications/global-risks-report-2024/digest/
- World Economic Forum. (2025, January 15). *Global risks report 2025*. https://www.weforum.org/publications/global-risks-report-2025/
- Xu, H. G., Costello, T. H., Schwartz, J. L., Niccolai, L. M., Pennycook, G., & Rand, D. (2025). Personalized dialogues with ai effectively address parents' concerns about HPV vaccination. *PsyArXiv*. https://osf.io/preprints/psyarxiv/gv52j_v1

- Yan, H. Y., Morrow, G., Yang, K.-C., & Wihbey, J. (2025, January 30). The origin of public concerns over AI supercharging misinformation in the 2024 U.S. presidential election. HKS Misinformation Review. https://misinforeview.hks.harvard. edu/article/the-origin-of-public-concerns-over-ai-supercharging-misinformation-in-the-2024-u-s-presidential-election/
- Zaller, J. R. (1992). *The nature and origins of mass opinion*. Cambridge University Press.
- Zimmerman, A., Janhonen, J., & Beer, E. (2024). Human/AI relationships: Challenges, downsides, and impacts on human/human relationships. *AI Ethics*, *4*, 1555–1567. https://doi.org/10.1007/s43681-023-00348-8
- Zagni, G., & Canetta, T. (2023, April 5). *Generative AI marks the beginning of a new era for disinformation*. European Digital Media Observatory. https://edmo.eu/2023/04/05/generativeai-marks-the-beginning-of-a-new-era-for-disinformation/
- Zarouali, B., Dobber, T., De Pauw, G., & de Vreese, C. (2020). Using a Personality-Profiling Algorithm to Investigate Political Microtargeting: Assessing the Persuasion Effects of Personality-Tailored Ads on Social Media. *Communication Research*, 49(8), 1066-1091. https://doi.org/10.1177/0093650220961965

About the Authors

FELIX M. SIMON is the Research Fellow in AI and News at the Reuters Institute for the Study of Journalism and a Research Associate at the Oxford Internet Institute (OII) at the University of Oxford. His research looks at the implications of a changing news and information environment for democratic discourse and the functioning of democracy. He regularly writes and comments on technology, media, and politics for various international outlets, including as a monthly columnist for Nikkei, and advises media organizations and companies on AI. Simon holds a DPhil and MSc from the University of Oxford's Internet Institute and a BA in Film and Media Studies from Goethe-University Frankfurt. He is affiliated with Columbia University's Tow Center for Digital Journalism and the Center for Information, Technology, and Public Life (CITAP) at the University of North Carolina at Chapel Hill.

SACHA ALTAY is a behavioral scientist working on misinformation social media and the news. He's currently a post-doctoral researcher at the University of Zurich within the political science department. He holds a PhD in cognitive science from the École Normale Supérieure in Paris.

About the Knight First Amendment Institute

The Knight First Amendment Institute at Columbia University defends the freedoms of speech and the press in the digital age through strategic litigation, research, policy advocacy, and public education. It promotes a system of free expression that is open and inclusive, that broadens and elevates public discourse, and that fosters creativity, accountability, and effective self-government.

knightcolumbia.org

