ANALYSIS 4/2024

AI-pocalypse Now?

Disinformation, AI, and the Super Election Year

Randolf Carr and Paula Köhler

October 2024

AI-enhanced disinformation was predicted to wreak havoc on elections around the world in 2024. However, the real negative effect of AI seems to have been limited. Several factors can explain why AI disinformation mostly fell flat, but they should not give rise to complacency. Technological and societal trends around AI indicate that greater disruptions to democratic processes are on the horizon.

The "super election year" 2024, with votes being cast in more than 60 countries accounting for over half the world's population, has coincided with a massive hype around generative artificial intelligence (GenAI).[1] Cutting-edge GenAI text, audio, images, and video can now seem indistinguishable from genuine, human-made content. Observers around the world have been sounding the alarm about how such convincing AI content could supercharge disinformation campaigns. Acute geopolitical tensions and intensified cooperation between authoritarian disinformation agents are heightening these fears.[2] So is polarization within democratic societies. Against this backdrop, the prospect of AI being used to influence voting outcomes prompted some to declare an "AI election year."[3]

With most of this year's votes cast, however, the dreaded "atomic bomb"[4] of AI disinformation has not (yet) detonated. While various examples of AI-generated content surfaced during elections, these appear not to have significantly impacted election campaigns or outcomes so far. The reason for the limited negative impact of AI is likely a combination of policy guardrails and industry norms around uses of AI, voter skepticism, and technological shortcomings. Yet, complacency is not called for. AI technology and disinformation tactics are continually evolving, and at least three trends indicate that "peak AI" for disrupting democratic processes is yet to come: More persuasive AI tools are being developed, AI content is becoming more pervasive, and public disengagement with political information is growing. Governments and tech companies therefore must continue developing policy and technological solutions, while civil society plays a critical role in strengthening resilience to disinformation in general.

Supercharged: The AI-Disinformation Nexus

Applying AI to disinformation operations can enhance their scale and sophistication. Disinformation campaigns can be dissected into three elements: actors, behaviors, and content. All three are impacted by advances in AI. Until recently, actors propagating disinformation needed access to and proficiency with software for tampering with image or audio material, as well as knowledge of their target audience's language. AI tools lower this skill and resource threshold, enabling more potential actors to generate more convincing deceptions.[5] At the same time, AI has revolutionized their behaviors for disseminating disinfor-

[1] GenAI describes a type of artificial intelligence that can create new text, images, video, audio, computer code, or other data based on patterns it identifies in existing data.

[2] Anne Applebaum, "The New Propaganda War," *The Atlantic*, May 6, 2024.

[3] Leah Feiger, "Welcome to the AI Election Year," *Wired*, May 31, 2024.

[4] Henry Foy, "Why Big Tech and Deepfakes Keep EU Election Guardians Up at Night," *Financial Times*, February 28, 2024.

[5] Josh A. Goldstein et al., "Generative Language Models and Automated Influence Operations: Emerging Threats and Potential Mitigations," Stanford: Stanford Internet Observatory, January 2023, https://perma.cc/V2HV-FQWC, 2.

mation: Prior to AI, actors commonly programmed vast networks of bot accounts whose automated behavior was often easy to detect. AI-powered bot armies are cheaper to run and less conspicuous. AI also helps embed the content posted by these bots. Seemingly authentic news websites, populated with innocuous AI-written articles, camouflage deceptive stories to lend them the veneer of credibility.[6] This AI-enabled "seeding" and "spreading" has become more difficult to detect than previous tactics.[7] Finally, disinformation content itself appears more convincing as AI can generate deceptively realistic visuals or audio, which are referred to as "deepfakes." A particularly insidious example is AI-generated non-consensual sexually explicit pictures and videos of women, which indeed make up an estimated 90 percent of all deepfakes.[8] Disinformation actors increasingly use these fakes to harass and push women and members of vulnerable groups out of political and public life.[9] Given the comprehensive potential for AI to enhance the potency of disinformation campaigns, policy-makers are thus rightly alarmed.

Artificial Hype? Use of AI-Generated Content in 2024 Elections

The "super election year" has put the issue of AI-enhanced disinformation operations in even sharper focus. Polling in the G7 countries, Brazil, China, India, and South Africa for the Munich Security Index 2024 showed public risk perception of "disinformation campaigns by enemies" and "artificial intelligence" rising steeply.[10] The World Economic Forum ranked "AI-generated misinformation and disinformation" as the second most likely risk to cause a "crisis on a global scale" in 2024.[11] In each of the 16 countries surveyed in a UNESCO poll in 2023, from Colombia to South Korea, majorities expressed worry about the use of AI to spread disinformation.[12] Consequently, many policy-makers, especially in Europe and the US, described in stark terms the election-disrupting threat of AI-enabled disinformation, with EU Commission Vice President Věra Jourová likening it to an "atomic bomb."[13] Tech sector leaders, too, have expressed concern that their AI could indeed "sway elections."[14]

However, through September 2024, AI-enabled tactics in disinformation campaigns have been both less prevalent and less impactful than widely expected. Perhaps the most concerning case came in late 2023 when an AI-generated audio surfaced of a supposed phone call between a journalist and the leader of the liberal Progressive Slovakia party discussing how to rig the upcoming election.[15] Progressive Slovakia suffered an upset loss at the polls days later, but what role the deepfake incident played in that is uncertain. In January 2024, up to 25,000 New Hampshire voters in the US presidential Democratic primary received "robocalls" in which an AI-generated voice of President Joseph Biden urged them not to vote.[16] In Taiwan's January election, deepfake videos promoted hoaxes such as a "secret history" of the outgoing leader, Tsai Ing-wen. Microsoft termed the China-based operation its first confirmed use of AI-generated material by a nation-state to influence a foreign election.[17] Both the US and Taiwanese cases heightened AI concerns, but the deceptions were quickly uncovered, and attempts to debunk them followed suit.[18]

> *AI-enabled disinformation campaigns have been both less prevalent and less impactful than widely expected.*

Elections in Pakistan, Indonesia, and India featured more prolific use of AI: GenAI was used to produce audio and video messages from candidates in languages they did not speak, from jailed candidates, and even from deceased political figures. For the most part, though, the generated material was not geared to deceive voters or otherwise malicious in nature.[19] Malicious AI-powered

[6] "Pro-Russia 'News' Sites Spew Incendiary US Election Falsehoods," *France24*, August 19, 2024.

[7] "Disinformation Is on the Rise. How Does It Work?," *The Economist*, May 1, 2024.

[8] Henry Ajder et al., "The State of Deepfakes: Landscape, Threats, and Impact," Amsterdam: Deeptrace Labs, September 2019, https://perma.cc/TQX5-NKQZ.

[9] Nina Jankowicz, "The Threat From Deepfakes Isn't Hypothetical. Women Feel It Every Day.," *The Washington Post*, March 25, 2021.

[10] Tobias Bunde, Sophie Eisentraut, and Leonard Schütte, "Munich Security Index 2024," in: Tobias Bunde/Sophie Eisentraut/Leonard Schütte (eds.), *Munich Security Report 2024: Lose-Lose?*, Munich: Munich Security Conference, February 2024, 26–45, https://doi.org/10.47342/BMQK9457, 32–33.

[11] World Economic Forum, "The Global Risks Report 2024," Geneva, January 2024, https://perma.cc/B7VT-FPEQ, 7.

[12] Lucia Mackenzie and Mark Scott, "How People View AI, Disinformation and Elections – in Charts," *Politico*, April 16, 2024.

[13] Foy, "Why Big Tech and Deepfakes Keep EU Election Guardians Up at Night," *Financial Times*.

[14] Dana Rao, "The Launch of the AI Elections Accord at the Munich Security Conference 2024," Munich: Munich Security Conference, February 16, 2024, https://perma.cc/Q7DS-CFTP.

[15] Rob Cameron and Ece Goksedef, "Slovakia Elections: Populist Party Wins Vote but Needs Allies for Coalition," *BBC*, October 1, 2023; Curt Devine, Donie O'Sullivan, and Sean Lyngaas, "A Fake Recording of a Candidate Saying He'd Rigged the Election Went Viral. Experts Say It's Only the Beginning," *CNN*, February 1, 2024.

[16] Max Matza, "Fake Biden Robocall Tells Voters to Skip New Hampshire Primary Election," *BBC*, January 23, 2024.

[17] Microsoft, "Same Targets, New Playbooks: East Asia Threat Actors Employ Unique Methods," Redmond, April 4, 2024, https://perma.cc/S8H5-6QA7, 6.

[18] Mareike Ohlberg, "Election Interference and Information Manipulation," Washington, DC: GMF, June 6, 2024, https://perma.cc/5SPZ-EKWF; Matza, "Fake Biden Robocall Tells Voters to Skip New Hampshire Primary Election."

[19] Yan Zhuang, "Imran Khan's 'Victory Speech' From Jail Shows A.I.'s Peril and Promise," *The New York Times*, February 11, 2024; Nilesh Christopher, "The Near Future of Deepfakes Just Got Way Clearer," *The Atlantic*, June 5, 2024; Heather Chen, "AI 'Resurrects' Long Dead Dictator in Murky New Era of Deepfake Electioneering," *CNN*, February 11, 2024; John Thornhill, "The Danger of Deepfakes Is Not What You Think," *Financial Times*, June 20, 2024.

tactics were even less prominent in the French, EU, and UK parliamentary elections in June and July. The few salient cases included French far-right campaigners using some AI-generated images, for instance depicting migrants arriving on France's shores. Similar right-wing GenAI imagery was used in the EU elections. In the UK, GenAI content only went "viral" in a handful of cases.[20]

It is difficult to quantify the effect of disinformation on any given election result. When comparing results to expected outcomes based on polling data, however, there are no clear signs that AI has swung an election to date.[21] Considering this and the limited number of known incidents, the impact of AI-enabled disinformation was, by all indications, a far cry from the nightmarish predictions for the 2024 election year.

Sigh of Relief: Four Reasons Why AI Disinformation Fell Flat

There are at least four factors that explain why AI tactics did not have the expected negative impact. First, the issue's high profile has led to action from both policymakers and private companies. In the EU, the Digital Services Act already regulates microtargeting and deceptive content. The new AI Act created additional information integrity requirements for GenAI. In the US, several states have bolstered existing campaigning laws with specific prohibitions on deepfakes.[22] Those involved in the New Hampshire robocalls face fines and criminal charges, which might deter imitators. Developers of some of the most popular GenAI tools have also worked to mitigate risks. At the Munich Security Conference 2024, major tech companies signed the AI Elections Accord, committing to combat the use of deceptive AI content.[23] Its goals include establishing technical safeguards, such as limiting what political or election-related content GenAI can put out and ensuring that AI-generated content is labeled or watermarked. Early evidence shows these efforts remain somewhat piecemeal,[24] but they have still likely raised the barrier to creating effective AI deceptions.

Second, campaigning industry norms and ethics may also be delaying the adoption of deceptive AI tactics. In the US, for instance, campaigners have been cautious about using GenAI due to potential reputational costs. Instead, campaigns have applied AI mostly for data analytics and targeting voters.[25]

Third, swinging an election with AI disinformation may be more difficult than expected due to citizens' information consumption and voting habits. Especially in polities with rising polarization, most voters hold firm voting preferences regardless of new information, real or fabricated. Most disinformation has also been shown to reach only a small fraction of voters, and most of those it does reach are already highly partisan, not undecided.[26] Publics around the world have also grown wary of AI-generated political content. In a US survey, over half of respondents familiar with ChatGPT said they would mistrust election-related information from the chatbot.[27] It is true that citizens' ability to discern online falsehoods, including AI-generated ones, is still relatively low and that they tend to overestimate their own ability.[28] Perhaps more importantly, though, citizens have generally become more skeptical toward any kind of information presented to them.[29]

Fourth, GenAI technology itself and the tactics of disinformation actors may just not have been ripe for the "AI election" moment in 2024. Across this year's elections so far, the lion's share of disinformation has been the conventional kind, such as misleading edits to genuine videos or images. Actors are already highly practiced at these methods. So, with AI content quality still having room for improvement, the utility of switching to AI tactics may not be high enough yet in many cases.[30]

[20] Valentin Châtelet, "Far-Right Parties Employed Generative AI Ahead of European Parliament Elections," Washington, DC: Digital Forensic Research Lab at the Atlantic Council, June 11, 2024, https://perma.cc/JQL4-GB48; Sam Stockwell, "AI Disinformation: Lessons from the UK's Election," Sydney: Australian Strategic Policy Institute, August 16, 2024, https://perma.cc/C9QT-ZR45; Mark Scott and Océane Herrero, "French Far-Right Parties Target Voters With AI Ahead of Vote," *Politico*, July 4, 2024.

[21] Sam Stockwell et al., "AI-Enabled Influence Operations: The Threat to the UK General Election," London: Centre for Emerging Technology and Security at the Alan Turing Institute, May 2024, https://perma.cc/6UXN-D7FW.

[22] Michelle M. Graham, "Deepfakes: Federal and State Regulation Aims to Curb a Growing Threat," *Reuters*, June 26, 2024.

[23] Munich Security Conference, "A Tech Accord to Combat Deceptive Use of AI in 2024 Elections," Munich, February 16, 2024, https://perma.cc/A4VT-2WXZ.

[24] Mark Scott, "Finally: Someone Used Generative AI in a Western Election: Digital Bridge Newsletter," *Politico*, July 4, 2024.

[25] Shane Goldmacher, Tiffany Hsu, and Steven Lee Meyers, "AI Promised to Upend the 2024 Campaign. It Hasn't Yet.," *The New York Times*, May 23, 2024; Mark Scott, "When Generative AI Gets Political: Digital Bridge Newsletter," *Politico*, June 6, 2024.

[26] Zeve Sanderson, Sol Messing, and Joshua A. Tucker, "Misunderstood Mechanics: How AI, TikTok, and the Liar's Dividend Might Affect the 2024 Elections," New York: Centre for Social Media and Politics, January 22, 2024, https://perma.cc/8DUB-2VD6.

[27] Colleen McClain, "Americans' Use of ChatGPT Is Ticking Up, but Few Trust Its Election Information," Washington, DC: Pew Research Center, March 26, 2024, https://perma.cc/68FV-FRZM.

[28] Molly Lesher, Hanna Pawelec, and Mercedes Fogarassy, "The OECD Truth Quest Survey: Methodology and Findings," Paris: OECD, OECD Digital Economy Papers 369, June 2024, https://doi.org/10.1787/92a94c0f-en, 26; Benjamin A. Lyons et al., "Overconfidence in News Judgments Is Associated With False News Susceptibility," *PNAS* 118:23 (2021), https://doi.org/10.1073/pnas.2019527118.

[29] "AI Will Change American Elections, but Not in the Obvious Way," *The Economist*, August 31, 2024.

[30] Mark Scott, "Microsoft Goes From Bad Boy to Top Cop in the Age of AI," *Politico*, May 7, 2024.

This year, the "atomic bomb" of AI disinformation may not have detonated, but the fuse has been lit. Thus, relief over a relatively uneventful "AI election year" must not lead to complacency.

No Time for Complacency: Three Trends for AI and the Information Space

There is no guarantee that the factors that dampened the impact of AI on this year's elections will hold going forward. Indeed, technological and societal developments already on the horizon are setting the stage for greater AI-fueled disruption. Three trends are especially concerning regarding AI and its potential impact on democracies.

> *The "atomic bomb" of AI disinformation may not have detonated, but the fuse has been lit.*

The first trend is the increasing persuasiveness of AI tools. GenAI may not have been sufficiently persuasive to revolutionize disinformation campaigns in 2024, but as tools rapidly become more sophisticated, their utility for malign actors will only increase. While GenAI disinformation has so far come chiefly in the form of audio and visual content, tools like chatbots, including so-called AI companions, may have a greater impact going forward. Such companion bots are designed to emulate human personalities and build relationships with their users. This is critical as disinformation becomes more persuasive when it comes from a trusted messenger.[31] Existing chatbots may be capable of generating convincing political arguments, yet they lack both the capability to personalize their messages and the trust of their audiences.[32] As companion bots become more lifelike and popular, AI-to-human conversation and trust may become normalized. As a result, the potential for such "AI friends" to act as vectors for misleading information will be high. Already, extremists are customizing conversational chatbots to, for instance, deny the holocaust.[33] Weaving disinformation into conversations with an AI that otherwise feel deceptively genuine and personal may be the dreaded game-changer for AI-powered manipulation.

The second trend is the growing pervasiveness of GenAI content and its impact on the political sphere. As GenAI content becomes ubiquitous across all areas of society, more political actors may drop their restraint in using it for campaigning. Doing so would contribute to blurring the line between legitimate political messaging and AI disinformation. This has already begun in the so-called Global South, where GenAI was used prolifically for "softfakes," video or audio ads attributing language skills or other positive traits to candidates. Even when clearly recognizable as AI-generated, softfakes have raised ethics concerns.[34] More importantly, they could be the start of a slippery slope: When candidate-approved softfakes or AI avatars become the norm, it will become more and more difficult to distinguish malicious deepfakes from legitimate political advertising or satirical content.[35] As AI-generated material that is near-indistinguishable from authentic content becomes widespread, so will public skepticism about what can be considered "real." Actors can make use of this skepticism to misrepresent real events as AI fabrications. This is termed the "liar's dividend."[36] For instance, in India, a candidate alleged that an audio clip in which he criticized his own party was a deepfake, even when independent fact-checkers found that the clip was authentic.[37] In reverse, the AI liar's dividend also opens new lines of attack. For example, after US President Biden announced that he would not stand as a candidate in the 2024 elections, false claims circulated that his Oval Office address and phone calls to Vice President Kamala Harris's campaign were deepfakes, casting doubts on his health.[38] The more authentic content and GenAI content exist indistinguishably side by side, the harder such claims are to debunk, and so, the liar's dividend grows. The ubiquity of political AI content will thus have confusing implications for citizens – but also for regulators, AI developers, and social media companies working on content moderation.

The third trend is AI contributing to public disengagement with political information and news in general. Pervasive AI content is already making it more difficult for citizens to sift through information online. The risk, then, may be less that citizens are deceived by AI content and more that they disengage from the information environment altogether. Recently, users across multiple social media platforms are complaining of being inundated with AI-generated spam not intended to mislead but merely to fish for attention.[39] Similarly, the number of counterfeit news websites populated exclusively by AI-written articles – some innocuous, some deceptive – has multiplied over the last year.[40] In this context, the

[31] American Psychological Association, "What Psychological Factors Make People Susceptible to Believe and Act on Misinformation?," Washington, DC, March 1, 2024, https://perma.cc/MT83-3ST6.

[32] Mark Scott, "UK's Digital Election Snoozefest: Digital Bridge Newsletter," *Politico*, June 27, 2024.

[33] David Gilbert, "Neo-Nazis Are All-In on AI," *Wired*, June 20, 2024.

[34] Rumman Chowdhury, "AI-Fuelled Election Campaigns Are Here – Where Are the Rules?," *Nature* 628 (2024), 237, https://doi.org/10.1038/d41586-024-00995-9.

[35] Christopher, "The Near Future of Deepfakes Just Got Way Clearer," *The Atlantic*.

[36] Bobby Chesney and Danielle Citron, "Deep Fakes: A Looming Challenenge for Privacy, Democracy, and National Security," *California Law Review* 107:6 (2019), 1753–1820, https://doi.org/10.15779/Z38RV0D15J, 1785.

[37] Josh A. Goldstein and Andrew Lohn, "Deepfakes, Elections, and Shrinking the Liar's Dividend," New York: Brennan Center for Justice, January 23, 2024, https://perma.cc/7KZN-DZR8.

[38] Monir Ghaedi, "Fact Check: Viral Video Claims Biden-Harris Call Made by AI," *Deutsche Welle*, July 24, 2024.

[39] Shannon Bond, "AI-Generated Spam Is Starting to Fill Social Media. Here's Why," *NPR*, May 14, 2024.

[40] "Disinformation Is on the Rise. How Does It Work?," *The Economist*.

material does not need to be deceptively realistic to have negative effects. Mediocre AI content is proliferating around the internet. As it does so, it can feed the very data pools used to train new AI models, which has sparked worries of a negative feedback loop leading to more mistake-prone GenAI tools.[41] The risk is that mediocre AI content increasingly diverts attention away from high-quality information sources, "muddying" the information space. This risk comes at a time when studies upon studies show that, around the world, citizens' trust in information sources like traditional media and public institutions is in decline and "news avoidance" is on the rise.[42] An intensifying barrage of untrustworthy or simply low-quality information online will thus only make citizens more cynical and mistrustful.[43] Given the democratic imperative of a well-informed electorate, citizens' disinterest in seeking out facts and high-quality information is a grave risk.

As technological and societal trends continue, the conditions for AI disinformation may be vastly different by the next major election cycle. This year was likely not the culmination of AI disrupting elections but just the beginning. Most importantly, these three trends illustrate how the impact of AI can go beyond elections to disrupting trust in democratic processes and the information environment itself.

Maintaining the Momentum Against AI-Enhanced Disinformation

The "super election year" has laser-focused international attention on AI's impact on democratic processes. While the nightmare scenarios have not played out so far, concerning trends for democracy remain. The growing persuasiveness of AI tools, the increasing pervasiveness of AI content, and the danger of broader public disengagement with political information require action by governments, the technology sector, and civil society. Relief about the 2024 elections therefore must not dampen existing momentum to implement regulatory and technological safeguards around AI and to strengthen societal resilience against disinformation of all kinds. Most of all, AI-proofing democracy will require all stakeholders to engage in constant cooperation, vigilance, and adaptation.

Key Points

1. AI exacerbates the threat of disinformation operations by lowering the entry threshold for actors and by giving them more potent tools for creating and disseminating deceptive content.

2. Contrary to predictions that AI disinformation would disrupt elections in 2024, however, AI has had a negligible impact so far and was used in mostly harmless ways.

3. Reasons for the negligible impact of AI disinformation include legislation and company self-regulation as well as campaigning industry norms around using AI, overall trends in information consumption and voting behavior, and technological shortcomings.

4. This is no reason for complacency, as technological and societal trends indicate that AI risks to democratic processes will intensify rather than subside. The most noteworthy trends include advances in persuasive AI tools, increasingly pervasive AI content, and growing public disengagement with political information. Governments, tech companies, and civil society must be vigilant towards these developments.

[41] Gonzalo Martínez et al., "Towards Understanding the Interplay of Generative Artificial Intelligence and the Internet," in: Fabio Cuzzolin/Maryam Sultana (eds.), *Epistemic Uncertainty in Artificial Intelligence*. Cham: Springer, 2024, 59–73, https://doi.org/10.1007/978-3-031-57963-9_5.

[42] OECD Publishing, "Building Trust and Reinforcing Democracy: Preparing the Ground for Government Action," Paris, OECD Public Governance Reviews, November 17, 2022, https://doi.org/10.1787/76972a4a-en; Nic Newman et al., "Digital News Report 2024," Oxford: Reuters Institute for the Study of Journalism, June 17, 2024, https://doi.org/10.60625/risj-vy6n-4v57.

[43] "AI Will Change American Elections, but Not in the Obvious Way," *The Economist*.

Authors
Randolf Carr is a Senior Policy Advisor at the Munich Security Conference.

Paula Köhler is a Policy Advisor at the Munich Security Conference.

Bibliographical Information
Randolf Carr and Paula Köhler, "AI-pocalypse Now? Disinformation, AI, and the Super Election Year," Munich: Munich Security Conference, Munich Security Analysis 4, October 2024, https://doi.org/10.47342/VPRS3682.

DOI: 10.47342/VPRS3682

Disclaimer
The Munich Security Conference does not express opinions of its own. The opinions expressed in this publication are the responsibility of the authors.

Contact
Stiftung Münchner Sicherheitskonferenz gGmbH
Karolinenplatz 3, 80333 München
www.securityconference.org
research@securityconference.org

Design: MetaDesign
Layout: Kathrin Strahl

Visit our app and social media channels:
www.linktr.ee/MunSecConf.