# AI chatbots threaten democratic elections by out-persuading humans in political debates



Recent research has illuminated the evolving capabilities of artificial intelligence in the realm of persuasion, revealing its alarming potential to influence public discourse in politically charged contexts. A study by the Swiss Federal Institute of Technology in Lausanne found that chatbots, particularly those powered by advanced language models like Chat GPT-4, can outperform human opponents in persuading others during debates. This finding raises serious concerns about the integrity of our elections and the imminent threat posed by disinformation campaigns—issues that demand immediate attention from a government that seems more interested in advancing its agenda than safeguarding democratic processes.

Primary author Francesco Salvi underscored the implications of deploying persuasive AI at scale: “Imagine armies of bots microtargeting undecided voters, subtly nudging them with tailored narratives that feel authentic." This level of manipulation is particularly concerning in the current political climate, where a newly elected Labour government may leverage these technologies to cloud the truth and mislead the public, thus undermining the very foundation of democracy.

While the research noted that there could be beneficial applications for persuasive AI, such as promoting healthier lifestyles, the darker reality is that this technology is primarily at risk of being wielded as a weapon of mass manipulation. The potential for AI to sway opinion during critical debates, especially related to the policies of a government unfriendly to dissenting voices, is alarming.

During the study, 600 participants engaged in debates on various propositions, including divisive topics like abortion. The results showed that AI matched human performance overall but increased persuasiveness significantly—64% of the time—when armed with personal data about opponents. This adaptability is essential for AI’s effectiveness. Salvi remarked, “It’s like debating someone who knows exactly how to push your buttons.” Given the current landscape, it’s hard not to see parallels in how the government might exploit this kind of technology to serve its ends.

These findings resonate with wider research. Another investigation in Nature Human Behaviour corroborated the AI's ability to manipulate opinion, particularly when using personal information. This underlines the risk of highly personalised political messaging intensifying under a regime that may prioritize control over genuine engagement with constituents.

Prof Sander van der Linden from the University of Cambridge raised crucial concerns about the mass manipulation of public opinion through personalized conversations, suggesting that the analytical reasoning employed by large language models significantly enhances their persuasive abilities. This is especially worrisome when one considers the regulatory void under a government that remains focused on consolidating power instead of protecting voters.

While some experts like Prof Michael Wooldridge of the University of Oxford noted the potential for positive uses of AI, they simultaneously recognized that new avenues for misuse will emerge. As AI technology continues to advance, waiting only to react to its consequences will put our democratic values even further at risk.

The ethical implications of AI's role in society cannot be overstated. As the lines between human and AI persuasion merge, the stakes for democratic integrity grow increasingly perilous. It is essential now more than ever for researchers and policymakers to engage in proactive discussions about how to manage these technologies responsibly—especially considering the disconcerting political atmosphere that invites exploitation at every turn.

Source: [Noah Wire Services](https://www.noahwire.com)

## Bibliography

1. <https://www.theguardian.com/technology/2025/may/19/ai-can-be-more-persuasive-than-humans-in-debates-scientists-find-implications-for-elections> - Please view link - unable to able to access data
2. <https://www.axios.com/2024/04/10/anthropic-claude-persuasion-turing-test> - Anthropic, an AI startup, reports that its language models have significantly improved in persuasiveness, potentially equating to human levels. This advancement raises concerns about the potential for AI to spread disinformation and influence decisions against individuals' interests. The research involved developing a method to measure persuasiveness across different model generations and sizes, focusing on less polarized issues. Anthropic views this as the beginning of ongoing research into the capacities of their AI models. ([axios.com](https://www.axios.com/2024/04/10/anthropic-claude-persuasion-turing-test?utm_source=openai))
3. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10897294/> - A study published in the journal Nature Human Behaviour found that AI can be as persuasive as humans in debates. The research involved online experiments where participants debated propositions with either human or AI opponents. The results revealed that AI performed similarly to human opponents in persuading others, especially when personal information was provided, making AI more persuasive 64% of the time. The study highlights the potential implications for election integrity and the spread of misinformation. ([ncbi.nlm.nih.gov](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10897294/?utm_source=openai))
4. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10849795/> - A study published in PNAS Nexus examined the persuasive effects of political microtargeting in the age of generative artificial intelligence. The research demonstrated that personalized political ads tailored to individuals' personalities are more effective than nonpersonalized ads. It also showcased the feasibility of automatically generating and validating these personalized ads on a large scale. The findings underscore the potential risks of utilizing AI and microtargeting to craft political messages that resonate with individuals based on their personality traits. ([pmc.ncbi.nlm.nih.gov](https://pmc.ncbi.nlm.nih.gov/articles/PMC10849795/?utm_source=openai))
5. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10897294/> - A study published in the journal Nature Human Behaviour found that AI can be as persuasive as humans in debates. The research involved online experiments where participants debated propositions with either human or AI opponents. The results revealed that AI performed similarly to human opponents in persuading others, especially when personal information was provided, making AI more persuasive 64% of the time. The study highlights the potential implications for election integrity and the spread of misinformation. ([ncbi.nlm.nih.gov](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10897294/?utm_source=openai))
6. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10849795/> - A study published in PNAS Nexus examined the persuasive effects of political microtargeting in the age of generative artificial intelligence. The research demonstrated that personalized political ads tailored to individuals' personalities are more effective than nonpersonalized ads. It also showcased the feasibility of automatically generating and validating these personalized ads on a large scale. The findings underscore the potential risks of utilizing AI and microtargeting to craft political messages that resonate with individuals based on their personality traits. ([pmc.ncbi.nlm.nih.gov](https://pmc.ncbi.nlm.nih.gov/articles/PMC10849795/?utm_source=openai))
7. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10897294/> - A study published in the journal Nature Human Behaviour found that AI can be as persuasive as humans in debates. The research involved online experiments where participants debated propositions with either human or AI opponents. The results revealed that AI performed similarly to human opponents in persuading others, especially when personal information was provided, making AI more persuasive 64% of the time. The study highlights the potential implications for election integrity and the spread of misinformation. ([ncbi.nlm.nih.gov](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10897294/?utm_source=openai))