Document 1 – X's AI chatbot spread voter misinformation – and election officials fought back

The Guardian, by Rachel Leingang, September 12, 2024

Soon after Joe Biden announced he was ending his bid for re-election, misinformation started spreading online about whether a new candidate could take the president's place.

Screenshots that claimed a new candidate could not be added to ballots in nine states moved quickly around Twitter, now X, racking up millions of views. The Minnesota secretary of state's office began getting requests for factchecks of these posts, which were flat-out wrong – ballot deadlines had not passed, giving Kamala Harris plenty of time to have her name added to ballots.

The source of the misinformation: Twitter's chatbot, Grok. When users asked the artificial intelligence tool whether a new candidate still had time to be added to ballots, Grok gave the incorrect answer.

Finding the source – and working to correct it – served as a test case of how election officials and artificial intelligence companies will interact during the 2024 presidential election in the US amid fears that AI could mislead or distract voters. And it showed the role Grok, specifically, could play in the election, as a chatbot with fewer guardrails to prevent the generating of more inflammatory content.

A group of secretaries of state and the organization that represents them, the National Association of Secretaries of State, contacted Grok and X to flag the misinformation. But the company didn't work to correct it immediately, instead giving the equivalent of a shoulder shrug […].

Thankfully, this wrong answer was relatively low-stakes: it would not have prevented people from casting a ballot. But the secretaries took a strong position quickly because of what could come next. […]

The effort worked. Grok now directs users to a different website, vote.gov, when asked about elections. […]

Musk has described Grok as an "anti-woke" chatbot that gives "spicy" answers often loaded with snark. Musk is "against centralized control to whatever degree he can possibly do that", said Lucas Hansen, co-founder of CivAI, a non-profit that warns of the dangers of AI. This philosophical belief puts Grok at a disadvantage for preventing misinformation, as does another feature of the tool: Grok brings in top tweets to inform its responses, which can affect its accuracy, Hansen said.

Grok requires a paid subscription, but holds the potential for widespread usage since it's built into a social media platform, Hansen said. And while it may give incorrect answers in chat, the images it creates can also further inflame partisan divides.

The images can be outlandish: a Nazi Mickey Mouse, Trump flying a plane into the World Trade Center, Harris in a communist uniform. One study by the Center for Countering Digital Hate claims Grok can make "convincing" images that could mislead people, citing images it prompted the bot to create of Harris doing drugs and Trump sick in bed, the Independent reported. The news outlet Al Jazeera wrote in a recent investigation that it was able to create "lifelike images" of Harris with a knife at a grocery store and Trump "shaking hands with white nationalists on the White House lawn".

"Now any random person can create something that's substantially more inflammatory than they previously could," Hansen said.

Document 2 – AI not a US election gamechanger yet, officials say

***Voanews.com*, by Jeff Seldin, September 11, 2024**

When the U.S. announced the seizure of 32 internet domains tied to Russian efforts to ply American voters with disinformation ahead of November's presidential election, prosecutors were quick to note the use of artificial intelligence, or AI.

The Russian operation, known as Doppelganger, drove internet and social media users to the fake news using a variety of methods, the charging documents said, including advertisements that were "in some cases created using artificial intelligence."

AI tools were also used to "generate content, including images and videos, for use in negative advertisements about U.S. politicians," the indictment added.

And Russia is far from alone in turning to AI in the hopes of swaying U.S. voters.

"The primary actors we've seen for election use of this are Iran and Russia, although as various private companies have noticed, China also has used artificial intelligence for spreading divisive narratives in the United States," according to a senior intelligence official, who spoke on the condition of anonymity in order to discuss sensitive information.

"What we've seen is artificial intelligence is used by foreign actors to make their content more quickly and convincingly tailor their synthetic content in both audio and video forms," the official added.

But other U.S. officials say the use of AI to spread misinformation and disinformation in the lead-up to the U.S. election has so far failed to live up to some of the more dire warnings about how deepfakes and other AI-generated material could shake-up the American political landscape.

"Generative AI is not going to fundamentally introduce new threats to this election cycle," according to Cait Conley, senior adviser to the director of the Cybersecurity and

Infrastructure Security Agency, the U.S. agency charged with overseeing election security.

“What we're seeing is consistent with what we expected to see,” Conley told VOA.

AI “is exacerbating existing threats, in both the cyber domain and the foreign malign influence operation-disinformation campaigns,” she said. But little of what has been put out to this point has shocked officials at CISA or the myriad state and local governments who run elections across the country.

“This threat vector is not new to them,” Conley said. “And they have taken the measures to ensure they're prepared to respond effectively.”

As an example, Conley pointed to the rash of robocalls that targeted New Hampshire citizens ahead of the state’s first in the nation primary in January, using fake audio of U.S. President Joe Biden to tell people to stay home and “save your vote.”

New Hampshire's attorney general quickly went public, calling the robocalls an apparent attempt to suppress votes and telling voters the incident was under investigation.

This past May, prosecutors indicted a Louisiana political consultant in connection with the scheme.

More recently, the alleged use of AI prompted a celebrity endorsement in the U.S. presidential race by pop star Taylor Swift.

“Recently I was made aware that AI of ‘me’ falsely endorsing Donald Trump’s presidential run was posted to his site,” Swift wrote in an Instagram social media post late Tuesday.

“It brought me to the conclusion that I need to be very transparent about my actual plans for this election as a voter,” she wrote, adding, “I will be casting my vote for Kamala Harris and Tim Walz.”

But experts and analysts say for all the attention AI is getting, the use of such technology in attacks and other influence operations has been limited.

“There's not a tremendous amount of it in the wild that's particularly successful right now, at least to my knowledge,” said Katie Gray, a senior partner at In-Q-Tel, the CIA's technology-focused, not-for-profit strategic investment firm.

“Most attackers are not using the most sophisticated methods to penetrate systems,” she said on September 4 at a cybersecurity summit in Washington.

Others suggest that at least for the moment, the fears surrounding AI have outpaced its usefulness by malicious actors.

‘We jump to the doomsday science fiction,” said Clint Watts, a former FBI special agent and counterterror consultant who heads up the Microsoft Threat Analysis Center (MTAC).

“But instead, what we're seeing is the number one challenge to all of this right now is access, just getting to the [AI] tools and accessing them,” he said, speaking like Gray at the cybersecurity summit.

Over the past 14 months, MTAC has logged hundreds of instances of AI use by China, Russia and Iran, Watts said. And analysts found that Moscow and Tehran, in particular, have struggled to get access to a fully AI toolbox.

The Russians “need to use their own tools from the start, rather than Western tools, because they're afraid they'll get knocked off those systems,” Watts said.

Iran is even further behind.

“They've tried different tools,” Watts said. “They just can't get access to most of them for the most part.”

U.S. adversaries also appear to be having difficulties with the underlying requirements to make AI effective.

“To do scaled AI operations is not cheap,” Watts said. “Some of the infrastructure and the resources of it [AI], the models, the data it needs to be trained [on] – very challenging at the moment.”

And Watts said until the products generated by AI get better, attempted deepfakes will likely have trouble resonating with the targeted audiences.

“Audiences have been remarkably brilliant about detecting deepfakes in crowds. The more you watch somebody, the more you realize a fake isn't quite right,” according to Watts. “The Russian actors that we've seen, all of them have tried deepfakes and they've moved back to bread and butter, small video manipulations.”

DM2

With the use of the two documents, answer the following question (about 250 words):

To what extent can AI be a threat in the 2024 US Elections?