Nation-states engage in US-focused influence operations ahead of US presidential election

MTAC REPORT

Date: April 17, 2024

Foreign malign influence in the 2024 US election began at a slower tempo than our Microsoft Threat Analysis Center (MTAC) team observed in both the 2016 and 2020 cycles. However, we are observing activity that may be informative of what might be coming.

Notably, the usual Russian election influence actors kicked into gear over the last 45 days. Russia's 2024 US election influence campaign at present employs a mix of themes from 2020 with a renewed focus on undermining US support for Ukraine.

The dominant conversation across global elections so far this year centers on malicious use of generative AI for propaganda and disinformation. We have created a process for assessing foreign manipulator use of AI to influence audiences. All three of the authoritarian actors reviewed in this report—Russia, Iran, and China—leveraged some form of generative AI to create content since last summer. We anticipate that election influence campaigns will include fakes—some will be deep, most shallow—and the simplest manipulations, not the most complex employment of AI, will likely be the pieces of content that have the most impact. That's our assessment of authoritarian nation-state use of generative AI to date, and we will dive into what we've seen thus far in more detail in this report.

This second election report from MTAC provides an update on what we've observed from Russia, Iran, and China and malicious use of AI since our November 2023 report "Protecting Election 2024 from foreign malign influence." This report supports and informs several of Microsoft's broader election defense initiatives to prevent influence and interference by authoritarian nation-states.

These broader initiatives include Microsoft's commitments in the Tech Accord to Combat Deceptive Use of AI in 2024 Elections led by Microsoft's Democracy Forward team.[1] MTAC's role in detecting and assessing nation-state activity during 2024's historic year of elections will support Microsoft's specific commitments in the Tech Accord to detect and respond to deceptive deepfakes and, more broadly, Microsoft's efforts to protect democratic institutions and information integrity.

Russian influence operations on US electorate focus on Ukraine headed into 2024

The deteriorated geopolitical relationship between the United States and Russia leaves the Kremlin with little to lose and much to gain by targeting the US 2024 presidential election. In doing so, Kremlin-backed actors attempt to influence American policy regarding the war in Ukraine, reduce social and political support to NATO, and ensnare the United States in domestic infighting to distract from the world stage. Russia's efforts thus far in 2024 are not

[1] https://blogs.microsoft.com/on-the-issues/2024/02/16/ai-deepfakes-elections-munich-tech-accord/

novel, but rather a continuation of a decade-long strategy to "win through the force of politics, rather than the politics of force," or active measures.[2] Messaging regarding Ukraine—via traditional media and social media—picked up steam over the last two months with a mix of covert and overt campaigns from at least 70 Russia-affiliated activity sets we track.

MTAC has identified several unique Russia-affiliated influence actors supporting this objective, each with its own methodology, content-generation capabilities, speed, and persistence. However, the most prolific of these actors are backed by or affiliated with the Russian Presidential Administration, highlighting the increasingly centralized nature of Russian influence campaigns, rather than relying principally on its intelligence services and the Internet Research Agency (known more commonly as the troll farm) as seen during the 2016 US presidential election. Each Russian actor has shown the capability and willingness to target English-speaking—and in some cases Spanish-speaking—audiences in the US, pushing social and political disinformation meant to portray Ukrainian President Volodymyr Zelensky as unethical and incompetent, Ukraine as a puppet or failed state, and any American aid to Ukraine as directly supporting a corrupt and conspiratorial regime.

The Russia-affiliated influence actor MTAC tracks as Storm-1516, as one example, has successfully laundered anti-Ukraine narratives into US audience spaces, with its content published in languages including English, Russian, French, Arabic, and Finnish. Storm-1516's method typically begins with a purported whistleblower or citizen journalist seeding the actor's disinformation on a purpose-built video channel, which is then covered by a seemingly unaffiliated network of managed or affiliated websites. These websites include sites based in the Middle East and Africa, as well as several English-language outlets such as DC Weekly, Miami Chronical, and The Intel Drop. Ultimately, after the narrative has circulated online for a series of days or weeks, US audiences repeat and repost this disinformation, likely unaware of its original source.

[2] https://www.npr.org/2018/04/25/586099619/the-russia-investigations-what-you-need-to-know-about-russian-active-measures

Figure 1: Storm-1516's process for laundering anti-Ukraine disinformation into US audience spaces.

Another prolific Russian influence threat actor focused on Ukraine, Storm-1099 (most widely known by its "Doppelganger" campaign[3]), has continued to use its elaborate network of forged media outlets as well as uniquely branded outlets and media projects to circulate anti-Ukraine propaganda. Some of Storm-1099's outlets explicitly focus on the US political sphere and the 2024 election, like "Election Watch" and "50 States of Lie." These outlets cover US political issues, promoting content on divisive social and geopolitical issues.

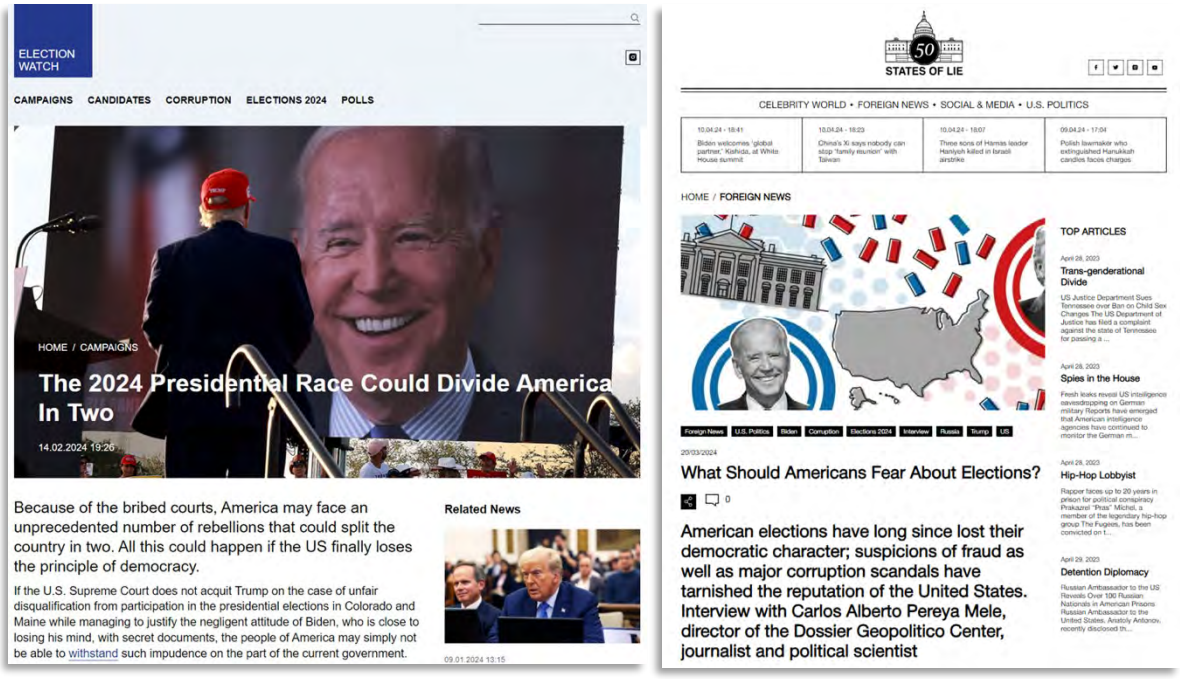

Figure 2: Storm-1099 US-focused websites posting election-related content.

[3] https://www.disinfo.eu/doppelganger-operation/

Russia is also again leveraging political influence ahead of November's contest via the latest iteration of the same campaign it launched during the 2020 US election cycle, dubbed "NABU Leaks," that targeted then-former Vice President Joe Biden. Ex-Ukrainian Parliamentarian and US-sanctioned Russian agent Andrei Derkach[4]—who spearheaded the now-sanctioned NABU Leaks campaign—reemerged on social media in early January 2024 for the first time since Russia's 2022 full-scale invasion of Ukraine. Derkach, in an interview conducted in Belarus and uploaded onto social media, propagated both old and new claims about US political figures, including President Biden. Ultimately, the aim of the NABU Leaks campaign, though focused on one presidential candidate, is about diminishing American support for Ukraine.

Figure 3: US-sanctioned Russian agent Andrei Derkach during the recent interview in which Derkach resurfaces the "NABU Leaks" campaign narratives.

Finally, hack-and-leak operations during elections separate Russian influence activities from those of other nations. In recent months, we've observed a notable uptick in cyber activity by Star Blizzard (formerly SEABORGIUM, also known as COLDRIVER, Callisto Group)—an FSB-affiliated threat actor that the UK government has accused of political interference.[5,6] Star Blizzard's latest campaigns focus on targeting western think tanks, and while the group's activity doesn't appear to have a direct connection to the 2024 election yet, Star Blizzard's current focus on US political figures and policy circles may be the first in a series of hacking campaigns meant to drive Kremlin outcomes headed into November. Microsoft Threat Intelligence will be closely monitoring Russian state-sponsored hacking groups' evolving activity, seeking to identify hacks designed to power influence operations headed into November.

China again targets American voters with covert influence operations

Like Russia, China's election-focused malign influence activity uses a multi-tiered strategy that aims to destabilize targeted countries by exploiting increasing polarization among the public and undermining faith in centuries-old democratic systems.[7] Tactically, China's influence

[4] https://home.treasury.gov/news/press-releases/sm1118

[5] https://www.gov.uk/government/news/uk-exposes-attempted-russian-cyber-interference-in-politics-and-democratic-processes

[6] https://www.microsoft.com/en-us/security/blog/2023/12/07/star-blizzard-increases-sophistication-and-evasion-in-ongoing-attacks/

[7] https://blogs.microsoft.com/on-the-issues/2023/09/07/digital-threats-cyberattacks-east-asia-china-north-korea/, https://blogs.microsoft.com/on-the-issues/2024/04/04/china-ai-influence-elections-mtac-cybersecurity/

operations seek to achieve these goals by capitalizing on existing sociopolitical divides, investing in target-country media outlets, leveraging cyber resources, and aligning its attacks with partisan interests to encourage organic circulation.

China's increasing use of AI in election-related influence campaigns is where it diverges from Russia. While Russia's use of AI continues to evolve in impact, People's Republic of China (PRC) and Chinese Communist Party (CCP)-linked actors leverage generative AI technologies to effectively create and enhance images, memes, and videos. One of the most prolific actors using AI content in influence operations is the CCP-linked actor Storm-1376 (known commonly as "Spamouflage"), which uses AI-generated content across a variety of themes to mislead audiences.

China's tactics have now seeped into influence operations targeting foreign elections and perceived Western adversaries. During the 2022 US midterm elections, for example, China sought to influence several midterm races involving members of both US political parties with the goal of countering candidates deemed to be "anti-China."[8] And, as we noted in our East Asia semi-annual report published on April 4, CCP-affiliated actors' use of AI included Taiwan's January 2024 presidential election.[9]

Figure 4: Sockpuppet accounts posting US-focused election content while posing as American voters.

[8] https://www.nytimes.com/2023/12/18/us/politics/election-interference-china-russia.html
[9] https://blogs.microsoft.com/on-the-issues/2024/04/04/china-ai-influence-elections-mtac-cybersecurity

Limited activity from Iran, a frequent late-game spoiler

Iran's election influence campaigns aim to weaken its adversaries by creating chaos, exacerbating existing political and social divisions, eroding trust in electoral processes or institutions, and undermining trust in the target country's leadership. Tehran's election influence strategy adopts a distinct approach: combining cyber and information operations for greater impact. During the 2020 US presidential election, Cotton Sandstorm—which is backed by the Islamic Revolutionary Guard Corps (IRGC)—conducted several cyber-enabled influence operations, including one leveraging voter details from a breached US voter registration database and giving the false appearance that fraudulent ballots had been cast. While past behavior suggests that Iran will likely launch acute cyber-enabled influence operations within weeks or months of Election Day, it is possible that Iran's planned goals and efforts directed at the United States will evolve with the ongoing conflict in the Middle East as a driver.

Generative AI in Election 2024: Risks remain but differ from the expected

The democratization of generative AI tools in late 2022 led to concerns that manipulators will use AI to create harmful propaganda and disinformation to potentially change the outcome of elections. MTAC's analytical teams have worked collaboratively with Microsoft's Responsible AI, Democracy Forward, and AI For Good Lab teams to identify, triage, and analyze nation-states' malicious use of generative AI in influence operations. We've logged incidents of nation-state actors using generative AI in campaigns across a spectrum of manipulation types. Our findings, thus far, suggest that the hypothesis positing that high-production, synthetic deepfake videos will create mass deception or broad-based confusion has not borne out. Rarely have nation-states' employments of generative AI-enabled content achieved much reach across social media, and in only a few cases have we seen any genuine audience deception from such content.

Instead, most of the incidents where we've observed audiences gravitate toward and share disinformation involve simple digital forgeries consistent with what influence actors over the last decade have regularly employed. For example, fake news stories with spoofed media logos embossed on them—a typical tactic of Russia-affiliated actors—garner several times more views and shares than any fully synthetic generative AI video we've observed and assessed.

The scenarios in which AI-generated or AI-enhanced content travels across social media at scale have considerable nuance. The following set of factors and conditions inform our team's assessment of generative AI risks as we head into a series of elections in 2024.

- **AI-enhanced, rather than AI-generated:** Fully synthetic deepfake videos of Russian President Vladimir Putin or Ukrainian President Volodymyr Zelensky are relatively routine in today's social media landscape. These videos, while using quite

sophisticated technology, are still not convincing and often rapidly debunked because the entirety of the video or near-entirety of the video is fabricated. Campaigns that mix both real and AI-generated content are more effective—a touch of AI-generated audio overlayed onto authentic video or integrating a piece of AI-generated content within a larger body of authentically produced content, for example—have been more convincing to audiences. These campaigns are also cheaper and simpler to create.

- Audio more impactful than video: Public concerns of generative AI employment have focused on the video medium, and particularly deepfake videos, but audio manipulations have consistently been more impactful on audience perceptions. Fake audio allegedly of politician Michal Šimečka and journalist Monika Tódová during the Slovak presidential election cycle is just one example of more convincing audio content.[10] Training data for generative AI audio is often more available for more people, requires less resourcing to create believable voices, less processing power, and remains harder to debunk without the context clues that AI-generated video can provide.
- Private setting over public setting: AI-generated audio has been more impactful in large part because of the setting in which audiences encounter it. Deepfake videos of world leaders have quickly been refuted by audiences who recognize oddities in the video or footage from the past. Collectively, crowds do well in sniffing out fakes on social media. Individuals independently assessing the veracity of media, however, are less capable. In private settings—during a phone call or on a direct-messaging application—inauthentic content can be difficult to assess, with no alternative opinions or subject-matter expertise by which to verify the authenticity of content.
- Times of crisis and breaking news: Soviet-era psychological warfare and today's Russian campaigns seize on calamitous messaging to push disinformation. Information consumers are more susceptible to deceptive content when scared or during fast-breaking events when the veracity of reported information may not yet be clear. Last summer, CCP-linked social media accounts published AI-generated images during the Maui wildfires in a coordinated disinformation campaign, as one example.[11] Those Advanced Persistent Manipulators (APMs) staffed with personnel equipped with generative AI tools will be well-positioned to deceive audiences headed into elections.

[10] https://www.wired.com/story/slovakias-election-deepfakes-show-ai-is-a-danger-to-democracy/
[11] https://www.nytimes.com/2023/09/11/us/politics/china-disinformation-ai.html

- Lesser-known impersonations rather than well-known impersonations: Audiences appear better at detecting manipulated content about individuals with whom they are more familiar. US voters have likely seen hundreds or even thousands of videos of 2024's candidates and thus will be more adept at identifying oddities in the inauthentic content. However, generative AI content regarding individuals or situations with whom an audience is less familiar—local election workers for example—or in languages or regions where an audience has less an understanding may be more impactful and pose a bigger risk to elections in the coming months.

Leading up to Election Day in the United States, MTAC will continue identifying and analyzing malicious generative AI use and will update our assessment incrementally as we expect Russia, Iran, and China will all increase the pace of influence and interference activity as November approaches. Of note: If there is a sophisticated deepfake launched to influence the election in November, the tool used to make the manipulation has likely not yet entered the marketplace. Video, audio, and image AI tools of increasing sophistication enter the market nearly every day.

The above assessment arises from what MTAC has observed thus far, but as both generative AI and geopolitical goals from Russia, Iran, and China evolve between now and November, risks to the 2024 election may shift with time.