FOREIGN AFFAIRS

NOVEMBER 19, 2025

AI Is Supercharging Disinformation Warfare

And America's Defenses Aren't Ready

JAMES P. RUBIN AND DARJAN VUJICA

AI Is Supercharging Disinformation Warfare

And America's Defenses Aren't Ready

JAMES P. RUBIN AND DARJAN VUJICA

In June, the secure Signal account of a European foreign minister pinged with a text message. The sender claimed to be U.S. Secretary of State Marco Rubio with an urgent request. A short time later, two other foreign ministers, a U.S. governor, and a member of Congress received the same message, this time accompanied by a sophisticated voice memo impersonating Rubio. Although the communication appeared to be authentic, its tone matching what would be expected from a senior official, it was actually a malicious forgery—a deepfake, engineered with artificial intelligence by unknown actors. Had the lie not been caught, the stunt had the potential to sow discord, compromise American diplomacy, or extract sensitive intelligence from Washington's foreign partners.

This was not the last disquieting example of AI enabling malign actors to conduct information warfare—the manipulation and distribution of information to gain an advantage over an adversary. In August, researchers at Vanderbilt University revealed that a Chinese tech firm, GoLaxy, had used AI to build data profiles of at least 117 sitting U.S. lawmakers and over 2,000 American public figures. The data could be used to construct plausible AI-generated personas that mimic those figures and craft messaging campaigns that appeal to the psychological traits of their followers. GoLaxy's goal, demonstrated in parallel campaigns in Hong Kong and Taiwan, was to build the capability to deliver millions of different, customized lies to millions of individuals at once.

Disinformation is not a new problem, but the introduction of AI has made it significantly easier for malicious actors to develop increasingly effective influence operations and to do so cheaply and at scale. In response, the U.S. government should be expanding and refining its tools for identifying and shutting down these campaigns. Instead, the Trump administration has been disarming, scaling back U.S. defenses against foreign disinformation and leaving the country woefully unprepared to handle AI-powered attacks. Unless the U.S. government reinvests in the institutions and expertise needed to counter information warfare, digital influence campaigns will progressively undermine public trust in democratic institutions, processes, and leadership—threatening to deliver American democracy a death by a thousand cuts.

INFORMATION AGE

For much of the modern era, many proponents of democracy have deemed the circulation of information to be purely a force for good. U.S. President Barack Obama famously articulated such a conviction in a speech to Chinese students in Shanghai in 2009, when he said that "the more freely information flows, the stronger the society becomes, because then citizens of countries around the world can hold their own governments accountable." Social media has accelerated the dissemination of information and made it easier for citizens to monitor, discuss, and raise awareness about government activities. But it has also undermined public trust in institutions and created online echo chambers through the promotion of personalized content and algorithms focused on engagement, limiting exposure to diverse viewpoints and deepening polarization among users.

Only in the last few years has the world finally recognized the urgency of the threats coming from the digital information domain. In a speech in

October, French President Emmanuel Macron drew a link between the exploitation of technology and democratic backsliding. Europe, he argued, has been "incredibly naive" to entrust its "democratic space to social networks that are controlled either by large American entrepreneurs or large Chinese companies." The political scientist Francis Fukuyama recently referred to this online public space as "an ecosystem that [rewards] sensationalism and disruptive content" and is shaped by the "unchecked power" of companies whose "interest of profit-maximization" leads to the unilateral amplification or suppression of particular voices—an outcome that goes against the core principles of democracy.

Advancements in AI have increasingly sharpened those threats to democracy. For the last half decade, countering foreign malign influence felt like tracking battleships in a game of naval warfare. U.S. adversaries such as China and Russia deployed large, state-controlled media outlets, including China's CGTN, for instance, and Russia's RT; clumsy fake social media profiles; and swarms of bots to push destabilizing narratives across the globe. Their methods were dangerous, but also blunt and easy to spot. Today, that era seems quaint. The disinformation battleships of old are still out there, but the rise of AI has opened the competition to a much wider array of combatants. Information warfare is now more akin to battle by autonomous drones—hyperpersonalized, relentlessly adaptive, and cheap enough for any actor to use against its adversaries. Foreign propaganda and disinformation campaigns are now engineered to seek out specific vulnerabilities—an individual's political leanings, social values, or even online shopping habits—and deliver targeted attacks designed to maximize the effects on their audiences' attitudes and behavior.

Propaganda campaigns have historically been constrained by the human labor required for content creation, translation, and target selection. AI removes those manpower demands, thus enabling information warfare to be waged at a speed and level of sophistication that many countries are not prepared to combat. Faced with an unstoppable onslaught of divisive political messaging, social cohesion could break down and government decision-making processes could become paralyzed. The digital information environment is now a theater of conflict in which domestic and foreign policy aims can be undermined by adversaries—all without requiring the attackers to leave the safety of their own territory.

BOTS WITHOUT BORDERS

The use of AI for intelligence gathering, disinformation campaigns, and malign influence operations is already spreading around the world. In El Salvador, for instance, President Nayib Bukele is fusing his sophisticated state propaganda apparatus with AI-powered tools, including bot networks. In addition to attracting foreign investment by putting the country's technological modernity on display, the use of AI bots is designed to help insulate the government from international criticism of its democratic backsliding by burying or rewriting narratives that allege human rights abuses.

AI is also being used to destabilize. OpenAI, the artificial intelligence company responsible for ChatGPT, recently announced that it had removed several ChatGPT accounts linked to Chinese actors. This covert influence operation, dubbed "Uncle Spam," used AI to create fake online personas and polarizing social media posts that deliberately argued multiple sides of contentious U.S. political issues, such as tariffs, with other social media users. The overall goal was to deepen political fractures within the United States. The component of "Uncle Spam" that was most corrosive to U.S. national security, however, was its attempt at intelligence gathering, which involved the use of AI tools to scrape and analyze vast amounts of personal data from platforms such as X (formerly Twitter) and Bluesky, including user profiles and follower lists belonging to American citizens. The Chinese-linked actors could use this information to refine their targeting methods, potentially giving Beijing an advantage in future rounds of information warfare.

Disinformation online can have consequences offline, too. In India, for example, a growing collection of AI-generated images and videos have spread hateful, anti-Muslim messaging, worsening existing interreligious tensions and fueling threats of psychological terror and physical violence against minority groups. According to a BBC report, in Sudan, where a civil war rages, AI voice cloning has been used on TikTok to impersonate former Sudanese leader Omar al-Bashir, who was ousted by the military in a 2019 coup and has not been seen in public for some time. Such a use of AI can degrade public trust in official sources of information and accelerate the breakdown of civil order amid an already brutal conflict.

Perhaps the most profound example of AI's power to disrupt occurred in Romania, where the 2024 presidential election was marred by foreign interference. A sweeping disinformation campaign, which Romanian intelligence services identified as linked to Russia, artificially boosted the online presence of Calin Georgescu, a far-right, pro-Russian fringe candidate. The operation included deepfakes, comments from tens of thousands of AI-powered bot accounts, and, according to authorities, payments to hundreds of influencers on social media platforms such as TikTok. The efficacy of the disinformation campaign was enough to put the legitimacy of the vote itself in question after Georgescu won in the first round of the election. Romania's Constitutional Court decided to annul the results, forcing a revote. The whole episode demonstrated that in some cases, AI-powered disinformation can not only threaten but also invalidate the fundamental processes of democracy.

STANDING DOWN

Even as the threat grows increasingly severe, the United States is more vulnerable to information warfare than ever before. In 2016, at the end of the Obama administration, the U.S. government started strengthening its ability to identify and counter foreign propaganda and disinformation—most notably with the establishment of the Global Engagement Center within the State Department. The GEC, along with other government offices focused on information warfare, used teams of geopolitical analysts and social media monitoring tools to unearth foreign influence campaigns. The State Department and the intelligence community also began studying adversarial tactics more closely and increased information sharing with foreign partners. But the U.S. government still struggled to keep up with advances in disinformation tactics.

The Biden administration made some progress. In 2023, the State Department, through the GEC, initiated a program to expose and disrupt Russia's information warfare campaigns in Africa and Latin America. The program employed a whole-of-government defense against disinformation: working with intelligence agencies to sanitize intelligence, stripping it of sensitive sources and methods to make it suitable for public use; with the Pentagon to assess the impact of information warfare on U.S. security; with the Treasury Department to impose sanctions; and with the White House to coordinate policy timing. In February 2024, a GEC-led effort resulted in the unearthing and dismantling of African Stream, an online media platform based in Kenya and secretly funded by Russia that spread anti-U.S. messaging, including stories that undermined confidence in American health programs. Perhaps most important, in September 2024, Secretary of State Antony Blinken announced that the United States would impose sanctions on Rossiya Segodnya, the parent company of the state-controlled television network RT. The sanctions were ordered after the State Department made public crucial intelligence that demonstrated how RT had become a clearinghouse for Russian covert information operations.

But the second Trump administration has cut or severely weakened the government offices responsible for identifying and countering foreign malign influence and disinformation campaigns. The GEC is among those offices, as are the Director of National Intelligence's Foreign Malign Influence Center, the FBI's Foreign Influence Task Force, and parts of the Cybersecurity and Infrastructure Security Agency, housed under the Department of Homeland Security. Eliminating this collection of offices means the U.S. government is no longer able to adequately identify, track, assess, and defend against adversaries in the information space.

The Trump administration's dismantling of these key agencies constitutes an irresponsible act of unilateral disarmament. After all, the bad actors are not going away. At the beginning of October, Ahmed Kaballo, the founder of African Stream, announced the launch of Sovereign Media, a self-described "anti-imperialist coalition" that promises to combat the "relentless censorship and algorithmic suppression" enacted by the "Western ruling elite." Sovereign Media's funding sources are unclear, but Kaballo is a longtime beneficiary of Russian largesse. Without the U.S. agencies that previously served as disinformation watchdogs, it is hard to know whether anyone in the Trump administration is taking a serious look at Sovereign Media or the many other foreign media outlets with connections to adversarial governments. Those actors, cumulatively, could do real harm to American interests as they flood the Internet with false narratives about the United States—especially as AI makes it increasingly difficult for citizens, both American and foreign, to separate false narratives from real ones.

ALL HANDS ON DECK

A perfect storm is brewing. U.S. adversaries are investing heavily in disinformation campaigns, AI advancements are ushering in a more dangerous form of conflict, and the second Trump administration has weakened the defenses that are meant to shield the United States and its partners from foreign malign influence.

There is no simple solution, but any serious U.S. defense against disinformation must entail both technological innovation and institutional restructuring. It should involve the United States' close allies and take a whole-of-government approach, one that includes a successor to the GEC and the reconstitution of other offices responsible for fighting disinformation. To aid in this effort, the Trump administration should issue a national security directive that unequivocally declares AI-amplified foreign malign influence a clear and present danger to the United States. This directive should mobilize the intelligence community to produce a new, comprehensive assessment of U.S. adversaries' disinformation capabilities, which would help focus future intelligence collection and targeting priorities on the most pressing threats. It should also establish a permanent interagency structure, led by the National Security Council, to ensure that tools available in different parts of the government, such as U.S. Cyber Command's offensive units and the Treasury Department's sanction mechanisms, are used in a coordinated fashion in the fight against foreign malign influence.

Defending against information warfare will also require partnership between the public and private sectors, organized by the White House Office of Science and Technology Policy. The creation of formal channels for collaboration with social media platforms, leading AI research labs, and cybersecurity firms would enable the U.S. government to share intelligence about particular threats, codevelop advanced technologies to help detect AI-generated content, and establish industry-wide best practices to counteract AI's magnification of disinformation. Through the White House's involvement, the fight against information warfare, now a niche policy concern, would become a central organizing principle of U.S. national defense.

Taking these steps is not meant to police free speech but rather to protect the right of American citizens to engage in dialogue that is unpolluted by foreign disinformation. With the 2026 U.S. midterm elections quickly approaching, the time to act is now. If the Trump administration fails to shore up the United States' defenses, the subtle and persistent influence campaigns deployed by its adversaries could undermine the democratic way of life that Americans hold dear.