

## Disrupting a covert Iranian influence operation

Archived: 2026-04-05 17:43:26 UTC

OpenAI is committed to preventing abuse and improving transparency around AI-generated content. This includes our work to detect and stop [covert influence operations](#) (IO), which try to manipulate public opinion or influence political outcomes while hiding the true identity or intentions of the actors behind them. This is especially important in the context of the many elections being held in 2024. We have expanded our work in this area throughout the year, including by leveraging our own AI models to better detect and understand abuse.

This week we identified and took down a cluster of ChatGPT accounts that were generating content for a covert Iranian influence operation identified as [Storm-2035\(opens in a new window\)](#). We have banned these accounts from using our services, and we continue to monitor for any further attempts to violate our policies. The operation used ChatGPT to generate content focused on a number of topics—including commentary on candidates on both sides in the U.S. presidential election – which it then shared via social media accounts and websites.

Similar to the covert influence operations we [reported in May](#), this operation does not appear to have achieved meaningful audience engagement. The majority of social media posts that we identified received few or no likes, shares, or comments. We similarly did not find indications of the web articles being shared across social media. Using Brookings' [Breakout Scale\(opens in a new window\)](#), which assesses the impact of covert IO on a scale from 1 (lowest) to 6 (highest), this operation was at the low end of Category 2 (activity on multiple platforms, but no evidence that real people picked up or widely shared their content). Our investigation benefited from information about the operation [published by Microsoft last week.\(opens in a new window\)](#)

Our investigation revealed that this operation used ChatGPT for two purposes: generating long-form articles and shorter social media comments. The first workstream produced articles on U.S. politics and global events, published on five websites that posed as both progressive and conservative news outlets. The second workstream created short comments in English and Spanish, which were posted on social media. We identified a dozen accounts on X and one on Instagram involved in this operation. Some of the X accounts posed as progressives, and others as conservatives. They generated some of these comments by asking our models to rewrite comments posted by other social media users.

The operation generated content about several topics: mainly, the conflict in Gaza, Israel's presence at the Olympic Games, and the U.S. presidential election—and to a lesser extent politics in Venezuela, the rights of Latinx communities in the U.S. (both in Spanish and English), and Scottish independence. They interspersed their political content with comments about fashion and beauty, possibly to appear more authentic or in an attempt to build a following.

Notwithstanding the lack of meaningful audience engagement resulting from this operation, we take seriously any efforts to use our services in foreign influence operations. Accordingly, as part of our work to support the wider community in disrupting this activity after removing the accounts from our services, we have shared threat intelligence with government, campaign, and industry stakeholders. OpenAI remains dedicated to uncovering and mitigating this type of abuse at scale by partnering with industry, civil society, and government, and by harnessing

the power of generative AI to be a force multiplier in our work. We will continue to publish findings like these to promote information-sharing and best practices.

---

Source: <https://openai.com/index/disrupting-a-covert-iranian-influence-operation/>