Twitter, Facebook ban fake users; some had AI-created photos

21 December 2019, by Mae Anderson

This July 9, 2019, file photo shows pedestrians walking across the street from the Twitter office building in San Francisco. Twitter says it has removed nearly 6,000 accounts it has deemed tied to a state-backed information operation in Saudi Arabia. (AP Photo/Jeff Chiu, File)

Twitter has identified and removed nearly 6,000 accounts that it said were part of a coordinated effort by Saudi government agencies and individuals to advance the country's geopolitical interests.

Separately, Facebook said it removed hundreds of Facebook accounts, groups and pages linked to inauthentic behavior from two separate groups, one originating in the country of Georgia and one in Vietnam, which targeted people both in Vietnam and in the U.S.

Facebook said some of the accounts used profile photos generated by artificial intelligence and masqueraded as Americans. It is one of the first such misinformation efforts to use material generated by AI.

Tech companies have stepped up efforts to tackle misinformation on their services ahead of next year's U.S. presidential elections. The efforts followed revelations that Russians bankrolled thousands of fake political ads during the 2016 elections to sow dissent among Americans.

Twitter's and Facebook's announcements underscore the fact that misinformation concerns aren't limited to the U.S. and Russia.

In a blog post Friday, Twitter said the removed Saudi accounts were amplifying messages favorable to Saudi authorities, mainly through "aggressive liking, retweeting and replying." While the majority of the content was in Arabic, Twitter said the tweets also amplified discussions about sanctions in Iran and appearances by Saudi government officials in Western media.

"Governments have started to launch influence campaigns the same ways commercial enterprises launch campaigns to sell detergent or cars," said James Ludes, a national defense expert who teaches international relations and public policy at Salve Regina University in Rhode Island.

He said the Russian efforts in 2016 showed it was possible to "actually change public attitudes through the targeted use of social media."

While the attempts to root out the campaigns may seem like a game of whack-a-mole, he said companies have at least shown progress in taking steps to identify and root out manipulation campaigns run by foreign powers.

Twitter began archiving tweets and media it deems to be associated with known state-backed information operations in 2018. It shut 200,000 Chinese accounts that targeted Hong Kong protests in August.

The 5,929 accounts removed and added to the archives are part of a larger group of 88,000

accounts engaged in "spammy behavior" across a wide range of topics. But Twitter isn't disclosing all of them because some might be legitimate accounts taken over through hacking.

The Twitter accounts were linked to a social media marketing firm in Saudi Arabia called Smaat that managed many government departments in Saudi Arabia. The accounts used third-party automated tools to amplify non-political content at high volumes. Twitter said that activity was used to mask the political maneuverings of the same accounts.

Samuel Woolley, a professor at the University of Texas at Austin who studies disinformation, said that while the Saudi campaign used basic manipulation techniques, including the use of likes and retweets to give the illusion of popularity, the campaign's size and scale were unusual. The existence of a thousands-strong army of Saudi accounts also show that social media companies still don't have a good solution, he said, despite the progress they have made at identifying state-backed accounts.

"It's really clear we have to do something about it," he said. "It can't just be after the fact. We have to get better about detecting in real time."

Messages left with Saudi officials in Riyadh, Saudi Arabia, and the country's embassy in Washington were not immediately returned.

The Saudi government has used different tactics to control speech and keep reformers and others from organizing, including employing troll armies to harass and intimidate users online. It has also arrested and imprisoned Twitter users.

In September, Twitter suspended the [account](#) of the crown prince's former top adviser, Saud al-Qahtani, who also served as director of the cyber security federation. As with Friday's announcement, Twitter said that account had violated the company's platform manipulation policy.

Last month, two former Twitter employees were charged with acting as agents of Saudi Arabia without registering with the U.S. government. The complaint details a coordinated effort by Saudi government officials to recruit employees at the social media giant to look up the private data of Twitter accounts, including email addresses linked to the accounts and internet protocol addresses that can give up a user's location.

In terms of Facebook's actions, Facebook said the Georgia group targeted domestic audiences and the Vietnam group focused mainly in the U.S., as well as Vietnamese-, Spanish- and Chinese-speaking audiences around the world.

The company said they created networks of accounts to mislead others about who they were and what they were doing. To evade detection, they used a combination of fake and real accounts of people in the U.S. to manage pages and groups, the company said.

"We are making progress rooting out this abuse, but as we've said before, it's an ongoing challenge," Nathaniel Gleicher, Facebook's head of security policy, said in a blog post.