Tech companies move to target terrorist propaganda online

December 5 2016, by Tami Abdollah

A Facebook log is displayed on the screen of an iPad, in this Wednesday, May 16, 2012 file photo taken in New York. Facebook, Microsoft, Twitter and YouTube announced a major shared effort Dec. 5, 2016, to more quickly identify the worst terrorist propaganda and prevent it from spreading online. The new program would create a database of unique digital "fingerprints" to help automatically identify videos or images the companies could remove. (AP Photo/James H. Collins, file)

Facebook, Microsoft, Twitter and YouTube are joining forces to more

quickly identify the worst terrorist propaganda and prevent it from spreading online.

The new program announced Monday would create a database of unique digital "fingerprints" to help automatically identify videos or images the companies could remove.

The move by the technology companies, which is expected to begin in early 2017, aims to assuage government concerns—and derail proposed new federal legislation—over social media content that is seen as increasingly driving terrorist recruitment and radicalization, while also balancing free-speech issues.

Technical details were being worked out, but Microsoft pioneered similar technology to detect, report and remove child pornography through such a database in 2009. Unlike those images, which are plainly illegal under U.S. law, questions about whether an image or video promotes terrorism can be more subjective, depending on national laws and the rules of a particular company's service.

Social media has increasingly become a tool for recruiting and radicalization by the Islamic State group and others. Its use by terror groups and supporters has added to the threat from so-called lone-wolf attacks and decreased the time from "flash to bang"—or radicalization to violence—with little or no time for law enforcement to follow evidentiary trails before an attack.

Under the new partnership, the companies promised to share among themselves "the most extreme and egregious terrorist images and videos we have removed from our services—content most likely to violate all our respective companies' content policies," according to a joint announcement Monday evening.

When such content is shared internally, the other participating companies will be notified and can use the [digital fingerprints](#) to quickly identify the same content on their own services to judge whether it violates their rules. If so, companies can delete the material and possibly disable the account, as appropriate.

Most social media services explicitly do not allow content that supports violent action or illegal activities. Twitter, for example, says users "may not promote violence against or directly attack or threaten other people on the basis of race, ethnicity, national origin, sexual orientation, gender, gender identity, religious affiliation, age, disability or disease."

"We really are going after the most obvious serious content that is shared online—that is, the kind of recruitment videos and beheading videos more likely to be against all our content policies," said Sally Aldous, a Facebook spokeswoman.

The new program caps a year of efforts to tamp down on social media's use by terrorist groups.

Lawmakers last year introduced legislation that would require social media companies to report any online terrorist activity they became aware of to law enforcement. The bill by Sens. Dianne Feinstein, D-Calif., and Richard Burr, R-N.C., was criticized for not defining "terrorist activity," which could have drowned government agencies in reports. The bill was opposed by the Internet Association, which represents 37 internet companies, including Facebook, Snapchat, Google, LinkedIn, Reddit, Twitter, Yahoo and others.

The bill came days after Syed Farook and his wife, Tashfeen Malik, went on a shooting attack in San Bernardino, California, killing 14 people and injuring 21 others. A Facebook post on Malik's page around the time of the attack included a pledge of allegiance to the leader of the

Islamic State group.

Facebook found the post—which was under an alias—the day after the attack. The company removed the profile from public view and informed law enforcement. Such a proactive effort had previously been uncommon.

Twitter moved toward partial automation in late 2015, using unspecified "proprietary spam-fighting tools" to find accounts that might be violating its terms of service and promoting terrorism. The material still required review by a team at Twitter before the accounts could be disabled.

"Since the middle of 2015, we have suspended more than 360,000 accounts for violating Twitter's policy on violent threats and the promotion of terrorism," said Sinead McSweeney, Twitter's vice president of public policy. "A large proportion of these accounts have been removed by technical means, including our proprietary spam-fighting tools."

Facebook has also used image-matching technology to compare images to ones it's already removed. The effort lets Facebook review images to avoid removing legitimate and protected uses, such as a photograph published by a news organization, a spokeswoman said.

Terrence McNeil of Ohio was charged in 2015 with soliciting the killings of U.S. service members over social media, including Tumblr, Facebook and Twitter. Federal prosecutors accused him of posting a series of photographs on his Facebook account to praise the death of a Jordanian pilot who was burned to death by the Islamic State group—showing him before, during and after his death, including an image of him engulfed in flames, according to the complaint.

In January, the White House dispatched top officials, including FBI

Director James Comey, Attorney General Loretta Lynch and National Security Agency Director Mike Rogers, to Silicon Valley to discuss the use of social media by violent extremist groups. Among the issues they discussed was how to use technology to help quickly identify terrorist content.

The four companies say they will be looking at involving additional companies in the future.

Citation: Tech companies move to target terrorist propaganda online (2016, December 5) retrieved 6 May 2024 from
https://phys.org/news/2016-12-tech-companies-terrorist-propaganda-online.html