FiveBy

Policy Advisory

Proposed Regulations Signal Need for Tech Companies to Rethink Content Monitoring

Executive Summary

A flood of proposed legislation to increase regulation over online content and recent contentious hearings on the subject indicate an appetite on Capitol Hill to hold tech companies accountable for disinformation and other harmful content hosted on their platforms. To mitigate possible liabilities, tech companies would be wise to increase content monitoring and analysis by linguistic, regional, cultural, and disinformation experts. Rising volumes of disinformation originating in Russia, Iran, and China underscore the necessity of specialized analytic expertise to supplement existing measures.

See Something, Say Something

On January 22, Senator Joe Machin (D-WV) proposed the See Something, Say Something Online Act of 2021 that would require tech companies to report suspicious content to support "criminal investigations and counterintelligence activities relating to international terrorism." The proposal would require tech platforms to monitor and report suspicious content through Suspicious Transmission Activity Reports (STARs)—similar to the Suspicious Activity Reports (SARs) financial institutions must file in accordance with the Bank Secrecy Act.

- The legislation would amend Section 230 of the Communications Decency Act to strip tech companies of their protections from being held legally liable for user activity if they do not file STARs to report suspicious social media content.
- The bill mandates the creation of an online system to file STARs administered by a Justice Department agency that would also be established under the legislation. Tech companies would have 30 days to submit a STAR and would need to keep the report on file for five years.
- The components of the STAR would identify the user's name and location, the date and nature of the user's post or other content, as well as time, origin, and destination and any relevant information or metadata related to the suspicious transmission.

More Possible Regulation on the Horizon

Skopos Labs projects that Manchin's bill has only a 4 percent chance of becoming law, but Congress and state legislatures clearly have an appetite for more regulation to counter social media disinformation and amend Section 230 to hold tech platforms at least partially accountable for what users post on their websites. The riots in Washington DC on January 6 amplified legislators' inclination to hold tech companies responsible for posts published on their sites, which will likely prompt platforms to mitigate their liability by increasing monitoring of user content or implementing controls that would restrain engagement, reducing users' ability or willingness to repost disinformation they see on the platforms.

- In February, Senator Mark Warner (D-VA) introduced the Safeguarding Against Fraud, Exploitation, Threats, Extremism and Consumer Harms (SAFE TECH) Act to clarify that Section 230 does not cover ads or paid content and increase liability for issues related to "civil rights, international human rights, antitrust and stalking, harassment, or intimidation." According to Skopos, this bill has a 30 percent chance of passage.

- Republicans have expressed discontent with social media companies targeting conservative content as disinformation. Section 230 does not obligate social media companies to conduct politically neutral content moderation. Legislation, such as the Platform Accountability and Consumer Transparency (PACT) Act, reintroduced by John Thune (R-ND) and Brian Schatz (D-HI) in March, would require tech companies to justify removing content and demonstrate that their moderation is neutral through biannual transparency reports. Analysts need disinformation expertise to moderate content and create these reports.

- Republicans in the Iowa Senate have advanced a proposal to outlaw tax breaks for and contracts with companies that censor free speech, and several other states are considering measures to allow social media companies to be sued for censorship.

Regional and Language Expertise Needed

Although some disinformation originates in the United States, most of the content comes from Russia, China, and Iran. Regional, cultural, and linguistic expertise would almost certainly allow tech platforms to recognize the origin of disinformation using language and regional analysis and help nip disinformation campaigns in the bud before they become widespread across US audiences and inflict reputational damage on tech company brands.

- Russia's disinformation strategy involves more than just social media accounts. An August 2020 State Department report found that Russia also spreads disinformation via state-funded global messaging, official government communications, proxy sources such as news websites that republish disinformation, ghostwriters that cite proxy sources to build credibility, and cyber-enabled disinformation which includes hack and release cyberattacks, cloned websites, and fake satellite imagery. Social media helps spread and amplify disinformation from these channels.

- China and Russia are both targeting Latin America with COVID disinformation campaigns, working to undermine confidence in the safety of the Pfizer vaccine and in the ability of the United States to manage the pandemic, according to SOUTHCOM Commander Craig Faller's Senate testimony in March. China is combining its disinformation campaigns with "mask and vaccine" diplomacy by distributing millions of Sinovac vaccines and masks in Latin America to bolster Beijing's political influence there. The Kremlin is targeting Latin America through coronavirus conspiracy theories and promotion of its Sputnik V vaccine on social media. Russia's goal is to build closer ties with Latin American countries, which could "help Russian military and security actors gain access to the United States' backyard," according to the Carnegie Endowment for International Peace.

- In November, the US Justice Department seized 92 domains linked to Iran's Revolutionary Guard Corps (IRGC) that were spreading disinformation about US foreign policy in Iran and the Middle East. At least one of the websites had social media presence on Twitter, Facebook, Instagram, and YouTube and claimed to operate in the United States, while listing an Iranian phone number. Similarly, In May 2020, Facebook dismantled eight Iran-linked networks involving more than 500 accounts that coordinated pro-Iran messaging campaigns targeted at Western voters to support Iran's geopolitical interests. Iran's PressTV network has also spread coronavirus conspiracy theories, along with criticism of the US maximum pressure campaign.

- Analysts can apply their regional and language expertise to uncover linguistic patterns and disinformation networks and typologies. An analyst with expertise in Russian historical disinformation techniques and strategies can expose Russian methodologies, such as the use of sleeper accounts that gain long term audience trust and spread false news. Linguistic, cultural, and regional expertise can also help identify disinformation methods such as those Russia employed before the 2016 US presidential election or during the 2016 "Lisa case" in which Russian media promoted a false narrative about a missing young Russian-German girl being raped by Arab immigrants to accuse Germany of tolerating child abuse, provoking protests by Russian Germans.

Conclusion

Although opposing political forces have different visions for social media reform, both are pushing for more regulation of online disinformation. Whether through STARs or other regulatory requirements, regulators have shown a significant appetite to track and mitigate disinformation that aims to spread online hate, stop the manipulation of information, and undermine US society's confidence in its government structures. The recent Global Trends report from the Office of the Director of National Intelligence highlights the growing power of social media that during the next 20 years will produce "content that could overtake expertise in shaping the political and social effects engendered by a hyperconnected information environment." Power increasingly will be wielded by the generators of content as well as the arbiters of who gets to see it, which legislators will almost certainly cite as justification for further regulation of online platforms. Analyzing data from social network communities that provides insight into the origination of disinformation, funding, and its proliferation will require deep understanding of language and financial transfer methodologies, as well as cultural and regional expertise.

FiveBy is a specialized risk intelligence services firm. We give you the insight you need to move faster and further with the confidence to transform your risks into opportunity. The opportunity to grow your profits, strengthen your brand, and exceed your customer expectations.

Our unique point of view brings together expertise spanning security, technology, data science, and business operations to connect your dots. By turning data into an enabler, FiveBy designs adaptable responses—whether to an ongoing incident or to implement preventive measures—tailored to your business needs and always with a human touch.